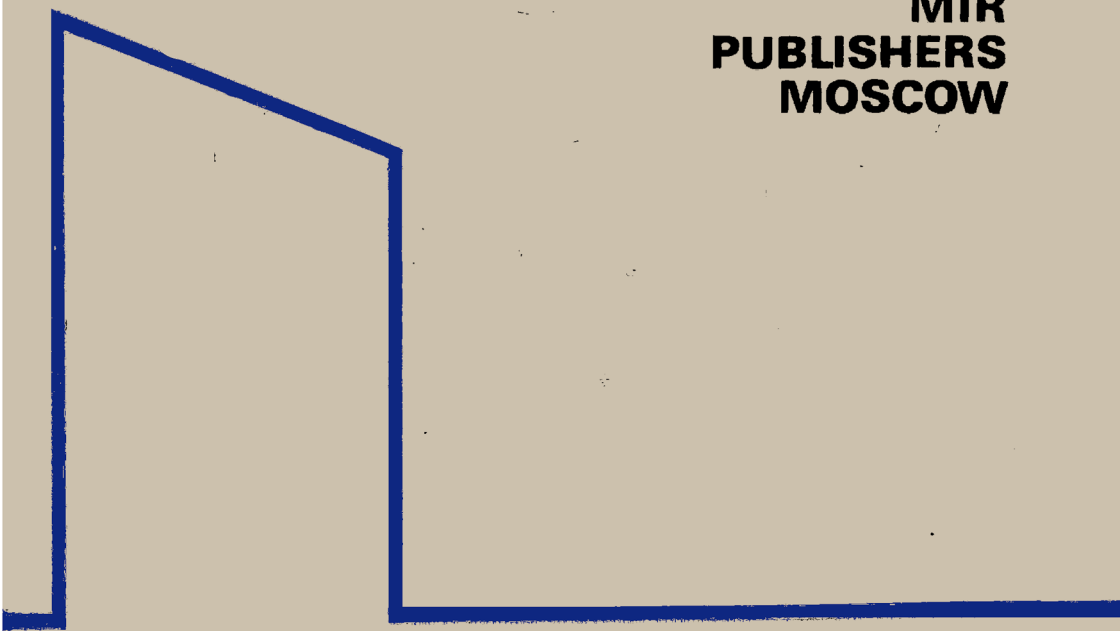
V.T. FROLKIN
AND
L.N. POPOV

PULSE CIRCUITS

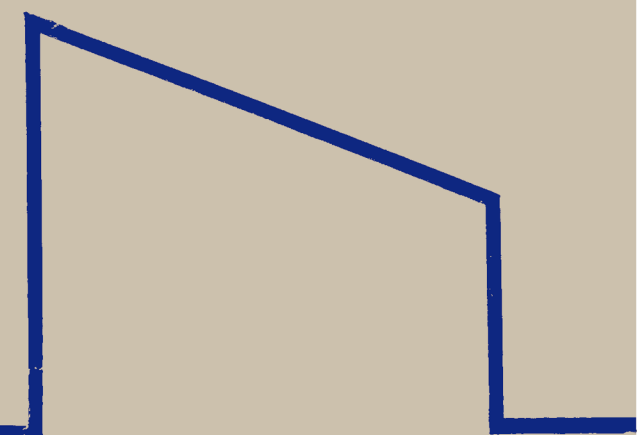
MIR
PUBLISHERS
MOSCOW



ABOUT THE BOOK

This book explains the analysis and operation of all the most frequently used pulse circuits. It covers solid-state circuits, including their IC forms. The text contains a large number of problems and review questions that will help the reader to get better insight into the subject-matter.

The authors start with pulse waveforms, move through *RS*-networks, to diode and transistor switching, logic gates, various forms of multivibrators, and flip-flops as building blocks for the more complex networks. The material also covers sawtooth voltage and current generators, relaxation oscillators, and their applications.



ИЗДАТЕЛЬСТВО
МИР

Фролкин В. Т., Попов Л. Н.

Импульсные устройства

Издательство Советское радио
Москва

V. T. Frolkin
and L. N. Popov

PULSE CIRCUITS

Translated from Russian
by
Boris V. Kuznetsov

Mir Publishers · Moscow

First published 1982
Revised from the 1980 third Russian edition

The Greek Alphabet

Α α	Alpha	Ι ι	Iota	Ρ ρ	Rho
Β β	Beta	Κ κ	Kappa	Σ σ	Sigma
Γ γ	Gamma	Λ λ	Lambda	Τ τ	Tau
Δ δ	Delta	Μ μ	Mu	Υ υ	Upsilon
Ε ε	Epsilon	Ν ν	Nu	Φ φ	Phi
Ζ ζ	Zeta	Ξ ξ	Xi	Χ χ	Chi
Η η	Eta	Ο ο	Omicron	Ψ ψ	Psi
Θ θ	Theta	Π π	Pi	Ω ω	Omega

The Russian Alphabet and Transliteration

А а	a	К к	k	Х х	kh
Б б	b	Л л	l	Ц ц	ts
В в	v	М м	m	Ч ч	ch
Г г	g	Н н	n	Ш ш	sh
Д д	d	О о	o	Щ щ	shch
Е е	e, ye	П п	p	ъ	''
Ё ё	e, ye	Р р	r	Ы ы	y
Ж ж	zh	С с	s	ь	'
З з	z	Т т	t	Э э	e
И и	i	У у	u	Ю ю	yu
Й й	y	Ф ф	f	Я я	ya

На английском языке

© Издательство «Советское радио», 1980 г.
© English translation, Mir Publishers, 1982

CONTENTS

1 GENERAL 9

1.1 Scope of the Course on Pulse Circuits 9
1.2 Analysis of Pulse Circuits 11

2 PULSE SHAPING BY LINEAR CIRCUITS 14

2.1 General 14
2.2 Pulse Response of a Series CR-Circuit Loaded by a Resistor 14
2.3 Pulse Response of a Series RC-Network Loaded by a Capacitor 21
2.4 Differentiation and Integration of Pulse Signals 24
2.5 Pulse Response of an RL-Network 28
2.6 Pulse Transformers 30
2.7 Delay Lines for Pulse Signals 37
Review Questions and Problems 46

3 SWITCHES 49

3.1 General 49
3.2 Diode Switches 50
3.3 Transistor Switches 63
3.4 Modifications of Transistor Switches 83
Review Questions and Problems 88

4 LIMITING AND CLAMPING CIRCUITS 92

4.1 General 92
4.2 Diode Limiters 94
4.3 Transistor Limiters 100
4.4 Switching Mode of Positive-Negative (Double-Peak) Limiters 106
4.5 Clamping Circuits 107
Review Questions and Problems 112

5 LOGIC GATES 114

5.1 Switching Functions 114
5.2 Basic Characteristics of Logic Gates 118
5.3 Common-Emitter Transistor Switch Logic Gates 127
5.4 Multistep Logic Gates 137

5.5 Current-Mode Logic Gates 146
5.6 MOSFET Logic Gates 153
Review Questions and Problems 158

6 BISTABLE CIRCUITS 161

6.1 General 161
6.2 Unsymmetrical Bistable Multivibrators 162
6.3 Symmetrical Bistable Multivibrators 167
6.4 Vacuum-Tube Bistable Circuits 177
6.5 Flip-Flops 178
6.6 Flip-Flops Made with Level Logic Gates 194
Review Questions and Problems 207

7 MULTIVIBRATORS 209

7.1 General 209
7.2 Discrete-Component Multivibrators 210
7.3 Modified (Integrated Circuit) Multivibrators 216
7.4 IC Logic Gate Multivibrators 222
7.5 Differential OPAMPs as Multivibrators 229
Review Questions and Problems 232

8 BLOCKING OSCILLATORS 234

8.1 General 234
8.2 Collector-Coupled Blocking Oscillator 234
8.3 Modified Blocking Oscillators 246
Review Questions and Problems 251

9 RAMP GENERATORS 252

9.1 General 252
9.2 CR Ramp Generators 257
9.3 Constant-Current Ramp Generator 262
9.4 Bootstrap and Miller Integrator Ramp Generators 266
Review Questions and Problems 283

10 SYNCHRONIZATION OF RELAXATION OSCILLATORS. PULSE FREQUENCY DIVISION 285

10.1 General 285
10.2 Synchronization of Free-Running Relaxation Oscillators 286
10.3 Pulse Rate Division by Triggered Relaxation Oscillators 289
Review Questions and Problems 291

11 GENERATION OF SAWTOOTH (RAMP) CURRENT 293

11.1 General 293
11.2 Transient Response of the Deflection System 294
11.3 Negative-Current-Feedback Sweep Generators 299
11.4 OPAMP Integrator Sweep Generators 307
11.5 Sawtooth (Ramp) Current Generators for Television 309
Review Questions and Problems 312

12 NEGATIVE-RESISTANCE PULSE CIRCUITS 314

12.1 General 314
12.2 Tunnel-Diode Pulse Circuits 314
12.3 Avalanche-Transistor Relaxation Oscillators 319
12.4 Thyristor Relaxation Oscillators 323
Review Questions and Problems 326

13 FUNCTIONAL PULSE AND DIGITAL CIRCUITS 328

13.1 General 328
13.2 Registers 328
13.3 Counters 348
13.4 Combinational Digital Functional Subunits 358
13.5 Analog Comparators 367
13.6 Pulse Time Gates 378
13.7 Variable Pulse Delay Circuits 380
Review Questions and Problems 382

BIBLIOGRAPHY 384

INDEX 387

1 GENERAL

1.1. SCOPE OF THE COURSE ON PULSE CIRCUITS

By *pulse circuits* we mean active and passive circuits intended to generate, shape, store and handle pulse and switching waveforms.

In pulsed operation, a circuit generates (shapes) or is acted upon by discontinuous signals appearing as pulses varying in waveshape. The term *waveshape* (or *waveform*) refers to the manner in which the associated voltage or current varies with time.

The basic pulse waveforms are *rectangular* (Fig. 1.1*a*), *trapezoidal* (Fig. 1.1*b*), *sawtooth* (*triangular*) (Fig. 1.1*c*) and *exponential* (Fig. 1.1*d*). They have no r.f. carrier and are called *video pulses.*

Figure 1.1 shows idealized pulse waveforms. They are only useful as linear models of zeroth approximation in pulse waveform analysis. Real pulses are not geometrically regular owing to the nonlinear behaviour of the tubes or semiconductor devices producing them and the presence of reactances in the associated circuits. Real rectangular (or square) pulses, most important in practice, have the waveshape shown in Fig. 1.1*e*. The major portion of the rise of a pulse is known as the *leading edge*, and the major portion of the fall of a pulse is referred to as the *trailing edge*. The portion between the leading and trailing edges is called the *top* of a pulse.

It is customary to define the interval of time required for the leading edge of a pulse to rise from 10% to 90% of its peak amplitude, V_m, as the *pulse rise time*, τ_r. Similarly, the interval of time required for the trailing edge of a pulse to fall from 90% to 10% of its peak amplitude is defined as the *pulse fall time*, τ_f. The time interval required for the instantaneous pulse amplitude to go from the 50-% point on the leading edge through the peak value and return to the 50-% point on the trailing edge is called *pulse duration* (*pulse length* or *pulse width*), τ_p. The pulse waveform shown in Fig. 1.1*e* has a *backswing* with an amplitude $V_{m,b}$. The illustration also shows damped sinusoidal oscillations (known as *ringing*) superimposed on the pulse top. The amplitude of the first maximum excursion is referred to as the *overshoot* of a pulse, ΔV.

A first-approximation linearized model of a rectangular pulse is shown in Fig. 1.1*f*. The straight-line segments *ab*, *bc* and *cd* approximate, respectively, the leading edge, top, and trailing edge of a real pulse. The rate of pulse voltage (or current) rise in Fig. 1.1*f*

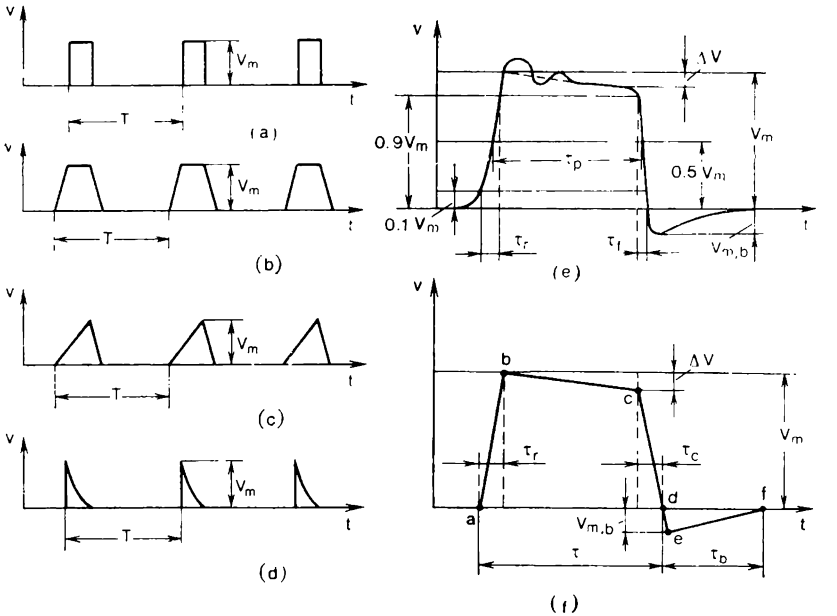


Fig. 1.1

is expressed in terms of the slope of the leading edge, $Sl = V_m/\tau_l$, and the rate of pulse top fall as the fractional *tilt*, *droop* or *sag*, $\delta V_m = \Delta V/V_m$.

The duration of real pulse signals ranges between broad limits: from a few nanoseconds to tens of seconds. Those most frequently used in computers, pulse communication, radar, TV and other fields of radio and electronics are microsecond pulses with durations from 10^{-7} to 10^{-1} s.

As often as not, the waveforms used in pulse circuitry are more elaborate in shape than those shown in Fig. 1.1. As an example, Fig. 1.2*a* shows an idealized voltage waveform appearing at the output of an *LR*-circuit driven by a sawtooth current. Figure 1.2*b* and *c* shows the current waveform in the deflection yoke and the voltage waveform in the beam-deflecting circuit of an alpha-numeric display CRT.

Intelligence transmitted by rectangular pulses can be embedded in some time function of the pulse: amplitude V_m, duration τ, and time position (repetition period) T.

According to the manner in which the intelligence-carrying time function of the pulse is modulated, existing pulse circuits may be classed into *analog* and *digital*. In analog pulse circuits, the intel-

ligence-carrying time function can, at least in principle, take on any values, however closely spaced, between the defined limits of an interval. In digital pulse circuits, the intelligence-carrying time function can take on only distinct values forming a finite set in an interval. The members of this set can be coded as digits in the binary, octonary, decimal or any other number system.

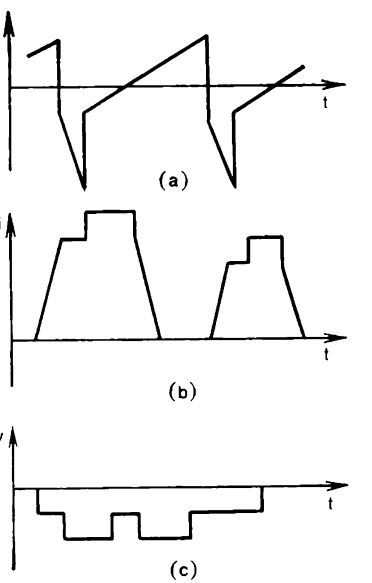


Fig. 1.2

Binary codes are preferable in many applications owing to their high noise immunity and also because they can readily be instrumented with switching circuits. Physically, a binary code uses level or pulse signals. In the level type of binary coding, an intelligence-bearing signal is a rectangular voltage waveform in which one level represents a logic 1, and another level a logic 0. In the pulse type of binary coding, the intelligence waveform is sampled at discrete clock intervals, and a logic 1 represents the presence of a sample pulse, and a logic 0 its absence.

Pulse working is implemented with linear circuits and electronic switches. The linear circuits are pulse amplifiers, *RCL*-circuits, pulse transformers, and delay lines. Of them, recourse is mostly had to pulse amplifiers and *RC*-circuits, because they can readily be microminiaturized.

Electronic switches are basically nonlinear devices, such as diodes, transistors, vacuum and gas-filled tubes, square *B*/*H* loop ferromagnetic cores, and negative-resistance semiconductor devices. As with linear circuits, preference is given to those of the switching circuits that can be built in microminiaturized form (diodes and transistors).

1.2. ANALYSIS OF PULSE CIRCUITS

Lumped-constant circuits are analyzed by Kirchhoff's laws equally applicable to both linear and nonlinear networks. Analysis of linear circuits in which the parameters of R, C and L elements and of controlled oscillators are independent of voltage and current relies on other simplifying rules, such as the superposition theorem, Thevenin's and Norton's theorems, etc.

In the pulse field, one is concerned with the analysis of the static and transient responses of circuits, sensitivity of intelligence-carrying

time functions or operating characteristics of pulse circuits towards destabilizing factors (noise, interference, climatic and mechanical factors, variations in the circuit parameters). The mathematical techniques used for the purpose are based on finite-difference or differential equations relating the unknown currents and voltages (circuit-state equations).

If a complex pulse circuit is to be analyzed and designed to the utmost of accuracy, it is essential to consider the nonlinear behaviour of its active and passive elements. As a result, the equations of state for such a circuit become nonlinear and have to be solved by numerical methods. One way to enhance the accuracy and speed of effort in such situations is through the use of computer-aided design techniques.

As the term itself implies, computer-aided design techniques are based on the use of computers and appropriate software (individual programs and program packages). The theoretical basis of computer-aided design is circuit analysis theory, and techniques of computational mathematics and programming. In the USSR, the combination of hardware and software intended for the above purpose has come to be known as a computer-aided design system (CADS).

The CADS is capable of handling a wide range of problems encountered in circuit and system design and engineering, and does so with an appreciable economic gain in terms of reduced time requirements, improved quality, and better coordination of the noncreative aspects automatically handled by a computer with the creative work that can only be done by the human designer manually.

Computer-aided circuit design on the basis of numerical methods removes most of the constraints otherwise imposed on the human designer in "manual" work by analytical methods. One of the crucial constraints is that the designer is forced to simplify as much as possible mathematical models for nonlinear elements by resort to the piecewise-linear approximation of their characteristics and to reduce the number of circuit elements in equivalent circuits. Since numerical methods readily supply solutions to practically any nonlinear equations likely to be encountered in circuit analysis and design, there is no need to simplify models for the circuit elements.

Another advantage of computer-aided design is that circuit engineering can now be extended to include large circuits containing hundreds or even thousands of elements and described, respectively, by hundreds and thousands of equations — a task unattainable in manual design and analysis.

Finally, computer-aided design provides a sufficiently accurate tool for handling statistical and circuit-optimization problems which, in manual analysis and design, are not taken up at all or treated very approximately, so that the results are only of qualitative value in all but the simplest cases.

The foregoing, however, is not meant to say that a computer can relieve the human designer of the necessity to evaluate circuit configurations on the basis of piecewise-linear models. As experience shows, cases are not infrequent in which the end results prove improbable because of the round-off errors inevitable in a computer because of an incorrect or an inaccurate statement of the original problem (for example, the designer may violate the constraints imposed on the range of values for model parameters adopted in a particular program). Such errors can be detected only if the designer knows an approximate solution close to the true one at least in terms of the order of magnitudes involved.

In determining the qualitative effect of pulse-circuit parameters on performance and an approximate quantitative estimate of the pulse-circuit behaviour, the designer ordinarily limits himself to rough calculations based on a piecewise-linear approximation of the *V-I* characteristics of active elements and linear equivalent circuits with time-switched parameters. In such cases, it is more convenient to use Laplace-transform calculus, and this is done in the present text.

2 PULSE SHAPING BY LINEAR CIRCUITS

2.1. GENERAL

In pulse work, pulses of a desired waveform are widely shaped by linear *RCL*-circuits. The simplest pulse-shaping circuit is a series *RC*-network. The output voltage can be picked off either its resistor or its capacitor—in either case the pulse-shaping response of the network will be different.

2.2. PULSE RESPONSE OF A SERIES CR-CIRCUIT LOADED BY A RESISTOR

Let the input voltage source have a zero internal resistance and the load resistance be infinitely large (Fig. 2.1). As the input voltage, we may use an ideal rectangular pulse and a ramp voltage. Such input signals provide sufficient insight into the pulse-shaping behaviour of the circuit and enable real rectangular pulses with edges of finite slope to be at least roughly approximated. The Laplace transform of the transmission gain of the circuit shown in Fig. 2.1 has the form

$$K(s) = s\tau_0/(1 + s\tau_0) \tag{2.1}$$

where $\tau_0 = RC$.

Single voltage pulse input. The initial change of voltage corresponding to the leading edge of an ideal rectangular pulse with an amplitude E (Fig. 2.2) will be passed by the capacitor on to the output without attenuation. Then, as the charge on the capacitor builds up, the circuit current and, as a consequence, the output volt-

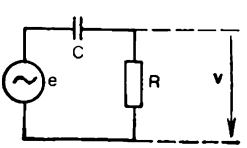


Fig. 2.1

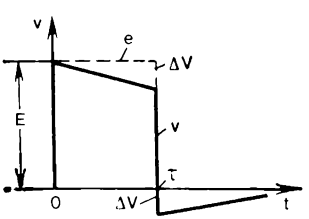


Fig. 2.2

age will decay exponentially:

$$v = E \exp(-t/\tau_0) \tag{2.2a}$$

At the instant when the pulse ceases ($t = \tau$), the output voltage will be $v(\tau) = E \exp(-\tau/\tau_0)$.

The trailing edge of the pulse is the negative change of voltage of amplitude E, appearing across the resistor. Thus, at the time $t = \tau$, the direction of current flow and the polarity of output voltage are suddenly reversed. The backswing of the output pulse, corresponding to the discharge of the capacitor during the interval $t \geqslant \tau$ varies as

$$v = -\Delta V \exp[-(t - \tau)/\tau_0] \tag{2.2b}$$

where

$$\Delta V = E - v(\tau) = E[1 - \exp(-\tau/\tau_0)]$$

Over the time from zero to τ, the charge on the capacitor is incremented by an amount

$$Q_{ch} = \int_0^{\tau} i\,dt = \frac{1}{R}\int_0^{\tau} v\,dt$$

After the pulse ceases, the charge is decremented by an amount

$$Q_{dis} = \int_{\tau}^{\infty} i\,dt = -\frac{1}{R}\int_{\tau}^{\infty} v\,dt$$

After all transients in the circuit die out, the initial charge on the capacitor remains unchanged, that is,

$$Q_{ch} - Q_{dis} = 0 \tag{2.3a}$$

Geometrically, Q_{ch} and Q_{dis} are represented by areas A_{ch} and A_{dis}, respectively, bounded by the output voltage waveform in the intervals from zero to τ and from τ to infinity. This leads us to a condition of practical importance— the equality of areas

$$A_{ch} = A_{dis} \tag{2.3b}$$

For an arbitrary waveform of the unipolar input signal, Eqs. (2.3) will hold if the time τ be taken as the instant when the circuit current and output voltage reverse in polarity. The condition defined by Eqs. (2.3) describes the loss of the d.c. component that can exist in the input signal.

When an *RC*-circuit is employed as a d.c. blocking network, it must distort the waveform of the transmitted pulse as little as practicable, that is, it must satisfy the following inequality:

$$\tau/\tau_0 \ll 1 \tag{2.4}$$

On expanding $\exp(-t/\tau_0)$ [see Eq. (2.2a)] into a power series and dropping all terms except the first one following the unity, we

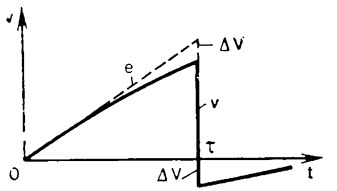


Fig. 2.3

obtain an expression describing the fractional droop of the output pulse (Fig. 2.2)

$$\delta V = \Delta V/E \approx \tau/\tau_0 \qquad (2.5)$$

Thus, for the maximum tilt (expressed as a percentage of the peak pulse amplitude) not to exceed 1 %, the time constant of the circuit must be at least 100 times the pulse duration.

When the circuit in question is driven by a single ramp pulse (Fig. 2.3) ($e = At$ for $0 \leqslant t \leqslant \tau$), in the interval from zero to τ the Laplace transform of the output voltage will be

$$v(s) = A\tau_0/s\,(1 + s\tau_0)$$

and the respective original time function will be

$$v(t) = A\tau_0\,[1 - \exp(-t/\tau_0)] \qquad (2.6)$$

At the time $t = \tau$, the input pulse ceases, and at other times $t \geqslant \tau$, the output pulse will vary as defined by Eq. (2.2b), but ΔV will be different:

$$\Delta V = e(\tau) - v(\tau) = A\tau - A\tau_0\,[1 - \exp(-\tau/\tau_0)] \qquad (2.7)$$

If Eq. (2.4) is satisfied, the distortion that a ramp pulse suffers on passing through a CR-network can be evaluated in terms of the ramp pulse tilt defined as

$$V_{\text{ramp}} = \Delta V/A\tau$$

(Fig. 2.3). To obtain a first approximation to V_{ramp}, the power series of the exponential functions in Eqs. (2.6) and (2.7) must now retain the quadratic terms. Then,

$$V_{\text{ramp}} \approx 0.5\tau/\tau_0 \qquad (2.8)$$

Pulse train input. If a pulse train with a repetition period T is to pass through a d.c. blocking network with little distortion, it is usually required to satisfy both the condition defined by Eq. (2.4) and the inequality (2.9)

$$(T - \tau)/\tau_0 \ll 1 \qquad (2.9)$$

As a consequence, the capacitor will have no time to recover its original charge between pulses. Now (Fig. 2.4*a*), the early pulses of the train bring about a transient rise in the residual charge on the capacitor, and some time later a state of dynamic equilibrium sets in. so that the increment in charge, ΔQ_{ch}, over the pulse duration becomes equal to the decrement in charge, ΔQ_{dis}, between pulses. As before, the values of ΔQ_{ch} and ΔQ_{dis} are proportional to the areas

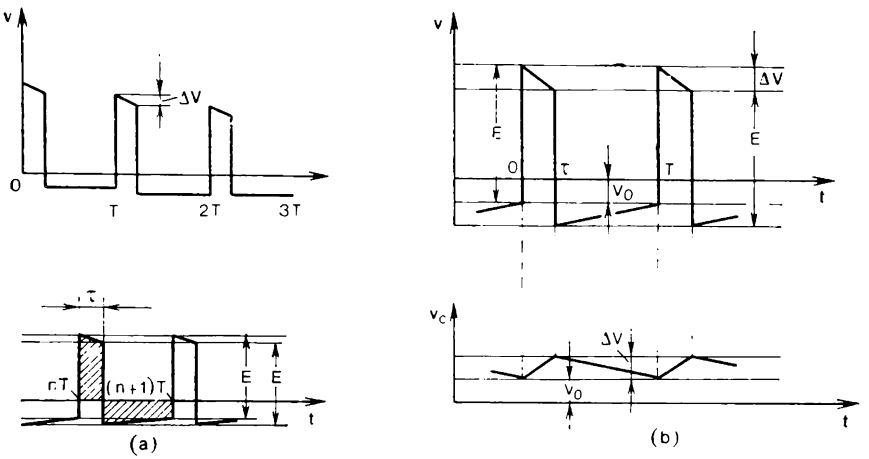


Fig. 2.4

bounded by the output voltage waveform over the intervals τ and T-τ (shown shaded in Fig. 2.4*a*).

In view of Eqs. (2.4) and (2.9), distortion in the waveshape of the output pulses can be ignored to a first approximation. Then the effect of the *RC*-network in question consists in shifting the initial level of the output pulses by V_0. If we agree to count the pulse duration from the instant when the circuit current and the output voltage reverse in polarity, the above conclusion can be extended to pulses of any arbitrary waveshape, $e(t)$. Over an nth interval, for the state of dynamic equilibrium we may write

$$v \approx \begin{cases} e - V_0 \text{ for } 0 \leqslant t \leqslant \tau \\ -V_0 \text{ for } \tau \leqslant t \leqslant T \end{cases}$$

From the equality of the shaded areas in Fig. 2.4*a*

$$\int_0^{\tau} (e - V_0)\, dt = \int_{\tau}^{T} V_0 dt$$

we get

$$V_0 = \frac{1}{T} \int_0^{\tau} e dt \equiv \frac{1}{T} \int_0^{T} e dt = \bar{e} \qquad (2.10a)$$

where $\bar{e}$ is the input voltage averaged over a period.

Changing the upper limit for the integral in Eq. (2.10a) from τ to T stresses that in the interval T-τ the applied voltage is zero. To sum up, transmission of a pulse train through a d.c. blocking network with low distortion entails a shift in the initial level, equal in a first approximation to the average value, $\bar{e}$, of the input signals

and independent of the circuit parameters. (Methods for changing the initial level of pulses at the output of an RC-network are examined in Chap. 4.)

Because the current charging the d.c. blocking capacitor in the state of dynamic equilibrium is smaller than the similar current for a single pulse, the tilt for a pulse train is smaller than the value given by Eqs. (2.5) and (2.8).

The value of δV for rectangular pulses can be found by reference to Fig. 2.4*b* which shows the output voltage v and the capacitor voltage v_C in the state of dynamic equilibrium. Locating the origin at point nT and using the generalized equation for an exponential function, $x(t) = x(\infty) - [x(\infty) - x(0)] \exp(-t/\tau_0)$, the exponential functions describing the charging and discharging of a capacitor can be written

$$v_C(t) = \begin{cases} E - (E - V_0) \exp(-t/\tau_0) \text{ for } 0 < t < \tau_0 \\ (V_0 + \Delta V) \exp[-(t-\tau)/\tau_0] \text{ for } \tau_0 < t < T \end{cases}$$

Using the boundary conditions, $v_C(\tau) = V_0 + \Delta V$ and $v_C(T) = V_0$, and limiting ourselves to the linear terms of the expansion, we get

$$\delta V = \Delta V/E \approx (\tau/\tau_0)(1 - \tau/T) \tag{2.10b}$$

Similarly, for a train of ramp pulses the tilt is

$$\delta V_{\text{ramp}} = \Delta V/A\tau \approx (\tau/2\tau_0)(1 - \tau/T) \tag{2.10c}$$

Pulse shortening. As often as not, the circuit of Fig. 2.1 serves to shorten pulses. The desired waveshape for the output signal produced by shortening an ideal rectangular pulse (Fig. 2.5) can be obtained by choosing the time constant for the circuit such that

$$\tau/\tau_0 \gg 1 \tag{2.11}$$

where τ is, as before, the duration of the input pulse.

The output bipolar pulses corresponding to the leading and trailing edges of the input pulse decay exponentially, as described by Eq. (2.2a). In an idealized case, the duration of these pulses, $\tau_{\text{out}} \approx 3\tau_0$, at 5% E may be shortened to any value, however small. In

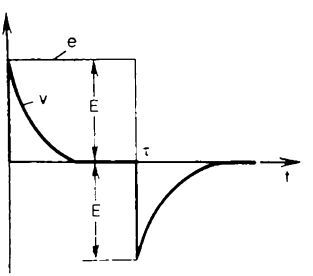


Fig. 2.5

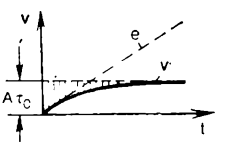


Fig. 2.6

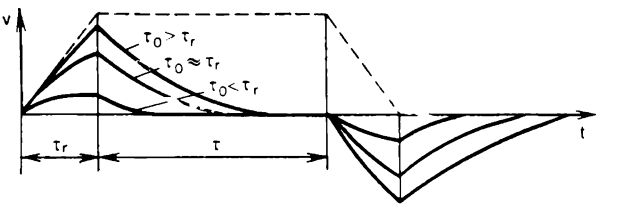


Fig. 2.7

practice, an unbounded decrease in the duration of output pulses with decreasing time constant of the circuit is prevented by the decrease in their amplitude, because the rise and fall times of the input pulses are finite and the previously neglected parasitic elements begin to be felt.

It is a relatively simple matter to evaluate the effect of the finite rise and fall times of the input pulse, if we assume that the voltages corresponding to the leading and trailing edges vary in a nearly linear fashion. To begin with, let us see how the output signal (Fig. 2.6) described by Eq. (2.6) varies under the conditions defined by Eq. (2.11). After the exponential term in Eq. (2.6) has decayed, the amplitude of the output voltage, $A\tau_0$, will remain constant and proportional to the rate of change of the input voltage. Thus, the circuit in question takes an approximate differential of the input signal, and the error of differentiation (defined as the difference in ordinate between the line $A\tau_0$ and the output voltage waveform) will increase as the time of differentiation is decreased. Ways and means of improving the accuracy of differentiation are discussed in Sec. 2.4.

Turning now to variations in the amplitude of a pulse with a finite-slope leading and trailing edge approximated by ramp waveforms, it is an easy matter to see that a decrease in the time constant of a pulse-shortening network will decrease the amplitude of the output voltage. If the condition defined by Eq. (2.11) is met, the effect of the exponential term will only be significant within the initial portion of the output pulse, and the amplitude of the latter will be directly proportional to the time constant (Fig. 2.7).

Effect of stray elements in a pulse-shortening RC-network. The principal stray elements that must be considered in pulse-shortening networks are the internal resistance R_s of the input signal source and the stray capacitive component, C_s, of the load impedance (Fig. 2.8). Because of these elements, the output voltage cannot change abruptly, so the shortened pulses may substantially differ from the idealized case discussed above in both waveform and amplitude.

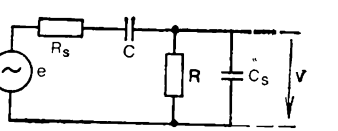


Fig. 2.8

If the input voltage e supplied by the source (Fig. 2.8) is a step input, $e = E$ at $t \geqslant 0$, that is, if $e(s) = E/s$, the Laplace transform of the output voltage will be

$$v(s) = E/\tau_s (s - s_1)(s - s_2) \tag{2.12a}$$

where $\tau_s = C_s R_s$, $\tau_0 = RC$, s_1 and s_2 are the roots of the characteristic equation

$$s^2 + s(1 + \gamma_R + \gamma_C)/\tau_s + 1/\tau_s\tau_0 = 0$$

For $\gamma_R = R_s/R$ and $\gamma_C = C_s/C$, these roots are

$$s_1 = -\frac{1}{2\tau_s}(1 + \gamma_R + \gamma_C)\left[1 - \sqrt{1 - \frac{4\gamma_R\gamma_C}{(1 + \gamma_R + \gamma_C)^2}}\right]$$

$$s_2 = -\frac{1}{2\tau_s}(1 + \gamma_R + \gamma_C)\left[1 + \sqrt{1 - \frac{4\gamma_R\gamma_C}{(1 + \gamma_R + \gamma_C)^2}}\right]$$

Noting that the fractions under the radical signs are smaller than unity, we can retain only the first two terms in the respective expansions. Then

$$\begin{aligned} s_1 &= -1/\tau_1 \approx -[\tau_0(1 + \gamma_R + \gamma_C)]^{-1} \\ s_2 &= -1/\tau_2 \approx -(1 + \gamma_R + \gamma_C)/\tau_s \end{aligned} \tag{2.12b}$$

The inverse Laplace transform (time function) of the output voltage (Fig. 2.9) is

$$v(t) = (E/\tau_s)\,[1/(s_1 - s_2)]\,[\exp(s_1 t) - \exp(s_2 t)] \tag{2.13}$$

where $|s_1| \ll |s_2|$.

As is seen from Fig. 2.9 where the exponential components of $v(t)$ are shown by a dashed line, the leading edge of the output pulse has at first a finite slope, $|dv/dt|_{t=0}$, governed by the ratio E/τ_s. By the time t_m determined primarily by the decay of the second exponential term in Eq. (2.13), the output pulse reaches a peak amplitude followed by a fall of voltage during which the controlling factor is the decay of the first exponential term in Eq. (2.13). It follows from Eq. (2.13) that a decrease in the circuit parameters of

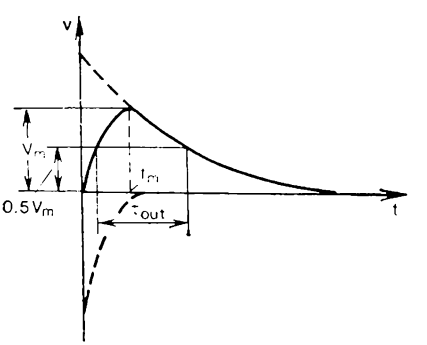


Fig. 2.9

a pulse-shortening *RC*-network will inevitable lead to a decrease in the amplitude of the output pulse.

An analysis (see, for example [2] and [31]) would show that when the duration of the output pulse and the amplitude of the input pulse are fixed in advance and the stray parameters remain unchanged, the parameters of the *RC*-network have certain optimal values, R_{opt} and C_{opt}, such that the amplitude of the output pulse is at a maximum.

2.3. PULSE RESPONSE OF A SERIES RC-NETWORK LOADED BY A CAPACITOR

In transform notation, the transmission gain of an *RC*-network loaded by a capacitor (Fig. 2.10) is

$$K(s) = (1 + s\tau_0)^{-1} \tag{2.14}$$

Step input. When a step input with an amplitude E is applied to the *RC*-network of Fig. 2.10, the output signal builds up exponentially:

$$v = E\,[1 - \exp(-t/\tau_0)] \tag{2.15}$$

The rise time of the output signal to 90%, 95% and 99% of E is $2.3\tau_0$, $3\tau_0$, and $4.6\tau_0$, respectively.

Rectangular pulse input. When the input signal is a rectangular pulse with a duration τ and an amplitude E, the leading edge of the output pulse is described by Eq. (2.15), and the trailing edge decays exponentially

$$v = v(\tau) \exp[-(t - \tau)/\tau_0] \tag{2.16}$$

where

$$v(\tau) = E\,[1 - \exp(-\tau/\tau_0)]$$

Pulse train input. The network of Fig. 2.10 is often a combination of the output resistance of an amplifying or a pulse-shaping stage and the stray capacitance ordinarily formed by three components, namely the output capacitance of the transistor, the wiring (circuit) capacitance, and the load capacitance. In the circumstances, the pulse-shaping action of the network is regarded as an undesirable effect on the waveshape of the input pulse. This effect takes the form of an exponential variation in the leading and trailing edges of the output voltage (curve *1* in Fig. 2.11). No generalized index has been adopted for this distortion yet. In practice,

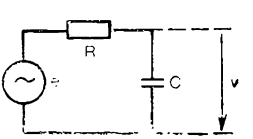


Fig. 2.10

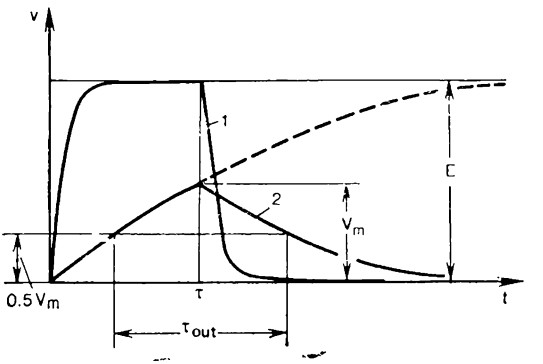


Fig. 2.11

the transmission of the input pulse may be taken as satisfactory if $\tau_r \leqslant 0.1\tau$. If the pulse rise time be taken as the time interval required for the output signal to become equal to 90%, 95% or 99% of E, then it is essential that τ_0 be less than or equal to 0.04τ, 0.03τ, or 0.02τ, respectively. If the pulse rise time be taken as the time interval required for the output signal to rise from 10% to 90% of E, equal to $2.2\tau_0$, a satisfactory transmission of an input pulse requires that τ_0 be less than or equal to 0.05τ. The same reasoning applies to the pulse fall time. The duration of the output pulse remains equal to that of the input pulse in all cases. Because the input and output terminals of the network shown in Fig. 2.10 are coupled conductively, the initial level of the input signal is transmitted without any change.

Pulse stretching. When the *RC*-network of Fig. 2.10 is used to stretch input pulses, the output pulse has the waveshape illustrated by curve *2* in Fig. 2.11. To a first approximation, the duration of the output pulse is given by

$$\tau_{out} \approx 0.5\tau + 0.7\tau_0 \qquad (2.17)$$

It is an easy matter to see that in the circuit in question an increase in the duration of the output pulse is accompanied by a fall in its amplitude, V_m, even under idealized conditions.

The principal parasitic element for a pulse-stretching *RC*-network is the load resistance R_L connected across the network capacitor C (Fig. 2.12*a*). The effect of this resistance can be evaluated

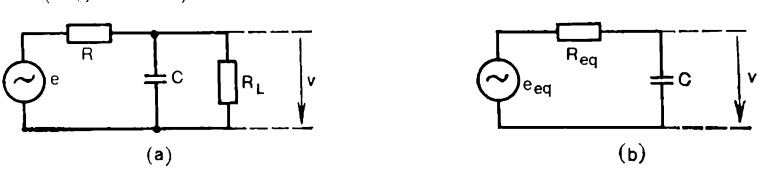


Fig. 2.12

by transforming the circuit of Fig. 2.12*a* into its equivalent circuit (Fig. 2.12*b*) on the basis of Thevenin's theorem. The emf e_{eq} of the equivalent voltage generator and the equivalent resistance R_{eq} used in the circuit of Fig. 2.12*b* are defined as

$$e_{eq} = eR_L/(R + R_L)$$

and

$$R_{eq} = RR_L/(R + R_L)$$

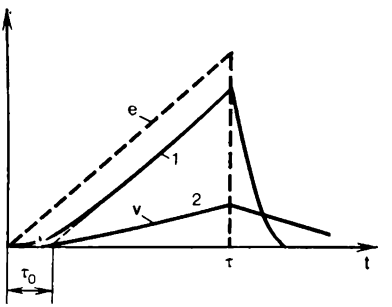


Fig. 2.13

Thus, the effect of the load resistance consists in reducing the amplitude and duration of the output signal.

Ramp voltage input. When the network of Fig. 2.10 is driven by a ramp voltage waveform, the output signal has a waveshape shown in Fig. 2.13 for two values of the relative pulse duration, τ/τ_0. Curve *1* illustrates the distortion brought about in the ramp input by a typical amplifying stage, when the capacitance of the *RC*-network is a parasitic one. Over the interval from 0 to τ, the output voltage can be described as

$$\begin{aligned} v(s) &= 1/s^2(1 + s\tau_0) \\ v(t) &= At - A\tau_0[1 - \exp(-t/\tau_0)] \end{aligned} \tag{2.18}$$

With a ramp input, the condition defined by Eq. (2.11) is usually satisfied. Therefore, the distortion produced in the output signal takes the form of a nonlinear initial portion and, after the exponential term in Eq. (2.18) has decayed, of a delay (shift) of the forward (rising) portion of the ramp waveform by τ_0. When the input ramp pulse ceases, variations in the output voltage, caused by the discharge of the capacitor C, are described by Eq. (2.16) where $v(\tau) \approx A(\tau - \tau_0)$.

Curve *2* applies when the ratio between the duration of the rising portion of a ramp input and the time constant of the network is as given by Eq. (2.4). On expanding into a power series the exponential term of Eq. (2.18) and retaining only the quadratic term, we obtain

$$v \approx 0.5At^2/\tau_0 \tag{2.19}$$

This voltage is proportional to the time integral of the input voltage. Thus, in the case on hand, the *RC*-network performs approximate integration.

2.4. DIFFERENTIATION AND INTEGRATION OF PULSE SIGNALS

The *RC*-networks examined above can be used for approximate differentiation and integration of input signals with respect to time. These operations are among the basic ones in analog computers and automatic process control systems.

Differentiation. The transfer function of a true differentiating element is

$$K_d(s) = s\tau_d \tag{2.20}$$

From a comparison of Eqs. (2.1) and (2.20), it is seen that the transfer function of an idealized *RC*-network (see Fig. 2.1) will approach $K_d(s)$ only as s tends to zero. Practically, differentiation is sufficiently accurate for the spectral components of the input signal which satisfy the following condition

$$\omega\tau_0 \ll 1 \tag{2.21}$$

It follows from Eq. (2.1) that when the above condition is satisfied, the amplitude of the output signal decreases.

Another factor standing in the way of reducing the time constant τ_0 in order to enhance the accuracy of differentiation is the effect of the stray parameters R_s and C_s which modify the transfer function to the value

$$K(s) = s/\tau_s (s - s_1)(s - s_2) \tag{2.22}$$

From a comparison of Eqs. (2.20), (2.1), and (2.22), it can be seen that a decrease in the circuit parameters of the *RC*-network is accompanied by an improvement in the accuracy of differentiation so long as the stray parameters can be neglected, that is, so long as γ_R and γ_C are much less than unity.

The basic approach to improving the accuracy of differentiation (as well as of integration), while keeping the amplitude constant, is to apply a compensating voltage; this matter will be taken up later.

Integration. The transfer function of a true integrating element is

$$K_1(s) = 1/s\tau_1. \tag{2.23}$$

Obviously, an *RC*-network loaded by a capacitor (see Fig. 2.10) will perform the operation of approximate integration only for the high-frequency spectral components of the signal for which

$$\omega\tau_0 \gg 1 \tag{2.24}$$

which, as with differentiation, leads to a reduction in the amplitude of the output signal.

Compensating voltage. It is an easy matter to see that the accuracy of differentiation and integration can be improved by minim-

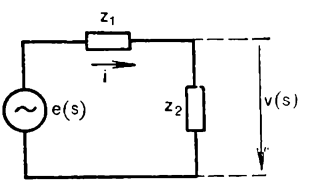


Fig. 2.14

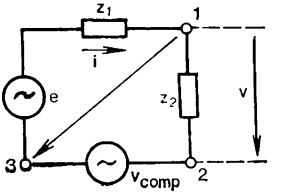


Fig. 2.15

izing the effect of the output signal on the current flowing in the network. To demonstrate, for the network of Fig. 2.14, we may write

$$\begin{aligned} v &= iZ_2 \\ i &= (e - v)/Z_1 \end{aligned} \tag{2.25}$$

If we assume that $v \ll e$, then

$$\begin{aligned} i &\approx e/Z_1 \\ v &\approx eZ_2/Z_1 \end{aligned} \tag{2.26}$$

Thus, if the impedances Z_1 and Z_2 are respectively a capacitor C and a resistor R, the network will perform the operation of differentiation. If, on the other hand, Z_1 R and Z_2 $1/sC$, the network will perform the operation of integration on the input signal.

Reduction in the amplitude of the output signal can be avoided by minimizing the dependence of current on voltage. This can be done by using a compensating voltage v_{comp}, as shown in Fig. 2.15. As a result, the current flowing around the circuit will be $i = (e - v + v_{comp})/Z_1$, and at $v = v_{comp}$ the current will be independent of the output voltage.

Generalized circuit of the operational amplifier. In the circuit of Fig. 2.15, the compensating voltage is supplied by a negative-feedback amplifier ordinarily referred to as an *operational amplifier* (*opamp*). In block-diagram form, it is shown in Fig. 2.16 where the notation for the signals and terminals is the same as in Fig. 2.15. Assuming that the opamp of Fig. 2.16 is an ideal voltage amplifier

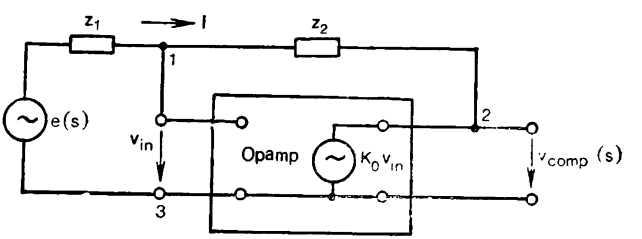


Fig. 2.16

(R_{in} is infinity and R_{out} is zero), then, using the principle of superposition, the input voltage v_{in} can be written

$$v_{in} = -v_{comp}/K_0 = (eZ_2 + v_{comp}Z_1)/(Z_1 + Z_2) \qquad (2.27)$$

Solving Eq. (2.27) for $K(s) = v_{comp}/e$ gives

$$K(s) = -\frac{Z_2}{Z_1(1+K_0)} \frac{K_0}{1+Z_2/(1+K_0)Z_1} \qquad (2.28)$$

The voltages v and v_{comp} are connected by a relation of the form

$$v = -v_{comp}(1 + K_0)/K_0 \qquad (2.29)$$

If the gain K_0 is chosen such that

$$K_0 \gg 1 \qquad (2.30a)$$

the equality of v and v_{comp} will be nearly complete. If the condition defined by Eq. (2.30a) is satisfied and the second term in the denominator of Eq. (2.28) is negligibly small so that

$$|Z_2/(1 + K_0)| \ll |Z_1| \qquad (2.30b)$$

then, by resolving the rational function of Eq. (2.28) into partial fractions and retaining only the first-order infinitesimals, we have

$$K(s) \approx -(1 - Z_2/K_0Z_1) Z_2/Z_1 \qquad (2.31)$$

It may be concluded from inspection of Eq. (2.31) that an opamp used under the specified idealized conditions can improve the accuracy of differentiation and integration K_0 times as compared with a simple RC-network, without reducing the amplitude of the output signal, because the opamp uses negative voltage feedback.

Basically, a practical opamp differs from the idealized circuit of Fig. 2.16 in that the input resistance is finite ($R_{in} \neq \infty$). Using the same approach as before and allowing for R_{in}, we obtain

$$K(s) = -\frac{K_0}{1+K_0} \frac{Z_1}{Z_2} \left[1 + \frac{Z_2/(1+K_0)}{Z_1 \| R_{in}}\right]^{-1} \qquad (2.32)$$

From a comparison of Eqs. (2.32) and (2.28), it is seen that the finiteness of R_{in} manifests itself in increasing the second (correction) term in the denominator of Eq. (2.32) which brings about a proportionate impairment in accuracy.

Integrated-circuit opamp of the differentiating type. A generalized circuit diagram of a differentiating opamp is shown in Fig. 2.17*a* where K_0 is the gain of the circuit in terms of the difference in input voltages at $Z_L = \infty$. Let K_0 be independent of the complex frequency variable s. This assumption will hold for slowly varying voltages and currents when the time lag (inertia) of the amplifier may be neglected. For fast signals, the dependence of K_0 on s may be

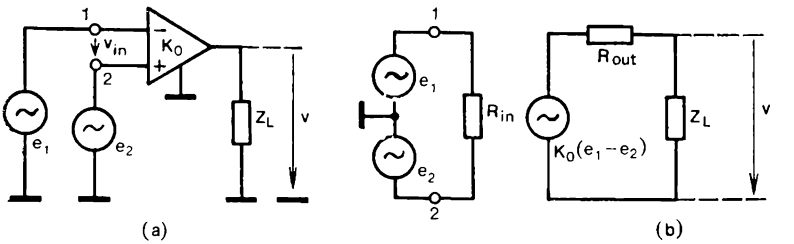


Fig. 2.17

a source of an additional error in differentiation and integration.

At $Z_L \neq \infty$, inspection of the quasi-static equivalent circuit of the amplifier input and output (Fig. 2.17*b*) gives

$$K_{0L} = v/(e_1 - e_2) = -K_0 Z_L/(R_{out} + Z_L) \tag{2.33}$$

As a rule, the inequality $Z_L \gg R_{out}$ holds for the main portion of the signal spectrum. Then, it may be deemed that $K_{0L} \approx K_0$. Terminal *1* is usually called the inverting input of an opamp, and terminal *2*, its noninverting (direct) input.

Let us derive the transfer function for a differentiating opamp (Fig. 2.18*a*), assuming that $R_{in} = \infty$ and $R_{out} = 0$. Invoking the principle of superposition, we obtain from the equivalent circuit of Fig. 2.18*b*

$$v_{in} = -v/K_0 = \frac{e_1 Z_2 + v Z_1}{Z_1 + Z_2} - e_2 \tag{2.34}$$

Hence,

$$v = -\frac{K_0}{1 + K_0} \frac{e_1 Z_2 - e_2 (Z_1 + Z_2)}{1 + Z_2/(1 + K_0) Z_1} \frac{1}{Z_1} \tag{2.35}$$

Grounding the noninverting input ($e_2 = 0$) in the circuit of Fig. 2.18*a* gives an operational amplifier. Now $K(s) = v(s)/e_1(s)$ is the same as defined by Eq. (2.28).

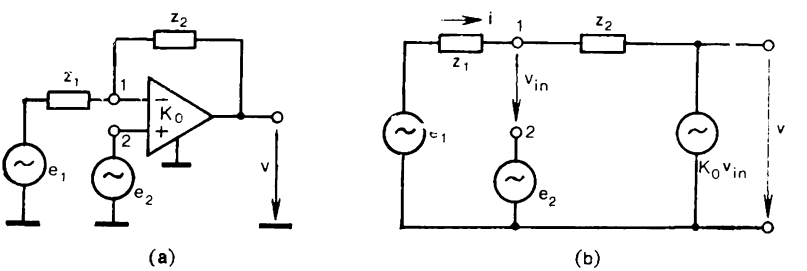


Fig. 2.18

It is of practical interest also to know the transfer function of an opamp referred to the noninverting input, when $e_1 = 0$:

$$K(s) = v(s)/e_2(s) = \frac{K_0(1 + Z_1/Z_2)}{1 + Z_1(1 + K_0)/Z_2} \tag{2.36}$$

Subject to the condition defined by Eq. (2.30a), the transfer function referred to the noninverting input does not, in a first approximation, depend on K_0:

$$K(s) \approx 1 + Z_2(s)/Z_1(s) \tag{2.37}$$

In particular, we can derive a circuit analogous to a highly efficient emitter follower, if we assume that $Z_1 = R$ and $Z_2 = 0$ or that $Z_1 = \infty$ and $Z_2 = R$.

When the differential input resistance is finite ($R_{in} \neq \infty$), the second (correction) term in the denominator of Eq. (2.35) increases as in the circuit of Fig. 2.16. Therefore, for the opamp to operate correctly (with the effect of K_0 minimized), it is essential, in addition to Eq. (2.30a), to satisfy also the condition defined by Eq. (2.38):

$$\left| \frac{Z_2}{R_{in} \| Z_1(1 + K_0)} \right| \ll 1 \tag{2.38}$$

The current trend in the USSR is to build operational amplifiers from integrated circuits [27, 42] and discrete elements Z_1 and Z_2. A typical IC differentiating opamp consists essentially of four stages, namely an input differential amplifier, a voltage amplifier, a d.c.-level shift circuit, and an output power amplifier [55]. The voltage gain of such amplifiers ranges from 10^4 to 10^5. The input stage may be based on bipolar and field-effect structures and have an input resistance of 10 kΩ to 10 MΩ and an output resistance of several tens or even hundreds of ohms.

2.5. PULSE RESPONSE OF AN RL-NETWORK

If in the previous RC-network we replace the C by an R and the R by an L, everything we have learned about the pulse-shaping action of RC-networks can be extended to the resultant RL-network. Unfortunately, such RL-networks are unsuitable for electrical differentiation and integration because real inductors have an ohmic series resistance, r, and a stray distributed shunt capacitance. For this reason, RL-networks are primarily used to shorten pulses in cases where the input pulses are relatively short. Then the inductors may be made small and without iron cores.

Figure 2.19*a* shows an inductor connected in the drain lead of a FET operating within the low-slope portion of the output V-I characteristics. From reference to the equivalent circuit of Fig. 2.19*b*, we find that in the circumstances the a.c. voltage v_- across the

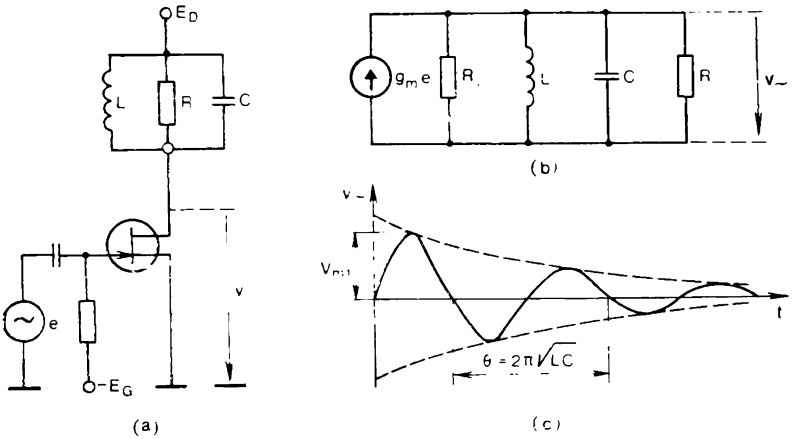


Fig. 2.19

inductor (with $r \approx 0$) is:

$$v_{\sim}(s) = \frac{g_m}{C} \frac{se(s)}{s^2 + s(CR_{eq})^{-1} + (CL)^{-1}} \tag{2.39}$$

where $R_{eq} = R_i \parallel R$.

For a step input voltage ($e = E$),

$$v_{\sim}(s) = (g_m E/C) \frac{1}{(s - s_1)(s - s_2)} \tag{2.40a}$$

$$v_{\sim}(t) = (g_m E/C) \frac{1}{s_1 - s_2} [\exp(s_1 t) - \exp(s_2 t)] \tag{2.40b}$$

where

$$s_1 = -\frac{1}{2} \frac{1}{CR_{eq}} (1 - \sqrt{\Delta})$$

$$s_2 = -\frac{1}{2} \frac{1}{CR_{eq}} (1 + \sqrt{\Delta}) \tag{2.41}$$

$$\Delta = 1 - (2R_{eq}/\rho)^2, \quad \rho = \sqrt{L/C}$$

Any one of several transient responses in the resonant circuit can be utilized to shape the output voltage, the choice being dependent on the value of R_{eq} which is in turn decided by the shunting resistance, R. The overdamped response is used to shorten pulses but seldom, but it frequently serves to attenuate output pulses of undesired polarity. For this purpose, a clipping diode is connected in parallel with the inductor.

As a rule, pulse-shortening networks are critically damped ($\Delta = 0$) or underdamped ($\Delta < 0$). In order to obtain a single output voltage pulse in the underdamped (or oscillatory) case, corresponding to the first half-cycle of a sine wave, a clipping diode must be placed

across the inductor. The diode must be connected in a polarity such that it is reverse-biased during the first half-cycle. When the voltage across the inductor reverses in polarity, the diode is forward-biased, and the network produces an overdamped response with a negligible amplitude because the impedance of the conducting diode is a small fraction of the characteristic impedance of the inductor. Sometimes, the inductor is replaced by a differentiating transformer [30] which, given a properly chosen transformation ratio, can substantially build up the amplitude of the output pulse.

Shock-excited resonant circuit. Given a step input and an oscillatory response ($\Delta < 0$), the sinusoidal oscillations generated in an inductor shunted by a capacitor are sometimes utilized to produce time marks—a sequence of pulses spaced accurately known and unvarying time intervals apart. Quite aptly, the circuit is called a *time-mark generator*. The accuracy of such a generator improves with an increase in the initial amplitude V_{m1} (Fig. 2.19*c*), a decrease in its attenuation from pulse to pulse, and an improvement in the stability of oscillation period θ. As a result, the transistor loaded into the inductor operates in the switching mode with a heavy initial current. In the no-signal condition, the transistor and inductor carry the quiescent current I_0, and the amplitude E of the controlling voltage change exceeds the cut-off voltage of the transistor. As a rule, no discrete shunting resistor need be included, and the resistance R refers to the equivalent resistance of the circuit (ordinarily called a shock-excited resonant circuit). This equivalent resistance combines the series resistance of the inductor and the equivalent load resistance. The basic component of the resonant-circuit capacitance C is that of the capacitor connected in parallel with the inductor to fix the frequency of the shock-excited sinusoidal oscillations. In the case of a time-mark generator, the amplitude of the shock-excited oscillations is usually stabilized by positive feedback [38].

2.6. PULSE TRANSFORMERS

Pulse transformers are intended to transmit or shape pulses with a duration from several nanoseconds to tens or even hundreds of microseconds. Above all, pulse transformers are required to transmit pulses with the least possible change in waveshape. They are widely used to control the amplitude, polarity and initial level of voltage and current pulses, to match and transform interstage impedances in pulse amplifiers, and to shorten pulses. A pulse transformer is a basic element of a blocking oscillator (see Chap. 8).

Transmission of voltage pulses. Let a step-input voltage e be impressed on the primary of a two-winding transformer at time t_0 (Fig. 2.20*a*). The transformer shown in the figure is an idealized one in the sense that it is assumed to have no stray parameters,

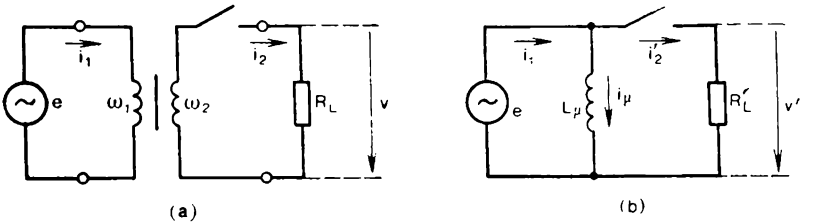


Fig. 2.20

such as series loss resistances, leakage inductances, shunt capacitances, eddy currents in the core, and variations in the permeability of the core material. The only difference of this transformer from an ideal one in the strict sense is that the primary inductance L_1 is finite, so the open-circuit magnetizing current is non-zero. However, the finite value of magnetizing current will not affect the waveshape of the output signal, because the internal impedance of the pulse generator is taken to be zero. An equivalent circuit for the transformer of Fig. 2.20*a*, referred to the primary winding, is shown in Fig. 2.20*b* where the magnetizing inductance L_μ and the parameters referred to the primary are

$$\begin{aligned} L_\mu &= \mu_a w_1^2 S_c / l_c \\ R'_L &= R_L / n^2 \\ n &= w_2/w_1, \quad i'_2 = n i_2, \quad v' = v/n \end{aligned} \tag{2.42}$$

where μ_a is the absolute permeability ($\mu_a = \mu\mu_0$), S_c is the cross-sectional area of the core, and l_c is the mean length of the core.

It follows from the equivalent circuit of Fig. 2.20*b* that a step input ($e = E$) brings about a linear rise in the magnetic induction B, the magnetic intensity H in the core, and the magnetizing current i_μ. As a consequence, on transmitting a pulse with an area $E\tau$, the core induction is incremented by

$$\Delta B = E\tau/(w_1 S_c) \tag{2.43}$$

In other words, in passing an ideal rectangular pulse of duration τ and amplitude E, the increments in the magnetic induction, magnetic intensity and magnetizing current are solely controlled by the transformer design and magnetic-circuit material and are independent of load. The only variable which depends on load is the primary current i_1 of the transformer. If no bound is placed on the generator current, then, in the absence of impedances between the input and output of the equivalent circuit in Fig. 2.20*b*, the magnetizing inductance L_μ will solely affect i_μ, and not the waveshape of the signal. In practice, the above assumptions do not hold, and the magnetizing current i_μ does affect the waveshape of the output signal (by reduc-

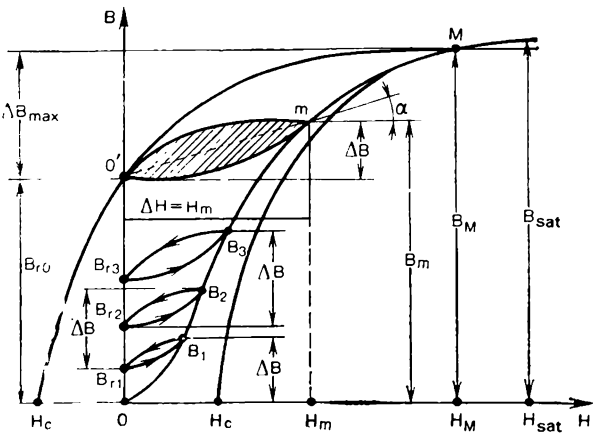


Fig. 2.21

ing its amplitude). As a consequence, L_μ ought not to be smaller than the required value.

In practice, pulse transformers depart from the idealized behaviour discussed above owing to magnetic hysteresis, eddy currents, leakage flux, and displacement current in the transformer windings.

Magnetic hysteresis. For the magnetic materials that go to make transformer cores, the magnetic induction is a nonlinear function of the magnetic intensity. Graphically, this relationship is depicted by what are called the normal magnetization curve and the hysteresis loop plotted for a limiting symmetrical cycle of magnetization and demagnetization when the current traversing the transformer primary varies at a relatively low rate (Fig. 2.21). The physical quantities usually stated in data sheets for transformer cores are the residual flux density or remanence B_{r0}, the coercive force H_c, the saturation flux density B_{sat} at some fixed value of H_{sat}. The point M is the cusp of the limiting hysteresis loop; beyond that point, the magnetization proceeds without hysteresis [31]. Let the initial state of magnetization in a core be designated by point 0 where the residual magnetization is zero. Then, as the transformer passes a sequence of pulses, the first pulse of area $E\tau$ will cause the representative point to move along the normal magnetization curve to point B_1 where the normal magnetization curve intersects a horizontal line, $B = \Delta B$, and the flux-density increment ΔB is given by Eq. (2.43).

When the first pulse ceases, the representative point moves to point B_{r1} along the descending arm B_1-B_{r1}. Arrival of the next pulse causes the representative point to move to B_2, first along the ascending arm of the curve, adjacent to the normal magnetization

curve, then along the normal magnetization curve, and the difference in flux density between B_2 and B_{r1} remains ΔB as it was before. When the second pulse ceases, the representative point moves to B_{r2} along the other descending arm. This build-up in remanence in response to arrival of every next pulse will continue until the representative point reaches 0′ corresponding to the remanence of the limiting cycle, B_{r0}. With arrival of every next pulse, the representative point will move between 0′ and m along the minor (unsymmetrical) hysteresis loop (shown shaded in Fig. 2.21) whose slope will be determined by the value of ΔB. Thus, the permeability averaged over a pulse, called the permeability on the limiting minor cycle, μ_Δ, will be given by (see Fig. 2.21):

$$\mu_\Delta = \tan \alpha = \Delta B/\Delta H \tag{2.44}$$

The value of μ_Δ is always smaller than the normal (quasi-static) permeability at point m ($\mu_a = B_m/H_m$), and the difference between them will increase with increasing remanence B_{r0}. To avoid any distortion in the waveshape of the output pulse, ΔB over the pulse should not exceed $\Delta B_{max} = B_M - B_{r0}$.

Thus, the core material for a pulse transformer should have the lowest possible remanence B_{r0}, a high saturation flux density B_{sat} and a high permeability μ_Δ.

At present, low-power pulse transformers predominantly use ferrite cores. Ferrites* have high permeability (about 1 500 gauss per oersted at $\Delta B < 2$ kilogauss [31]), a good temperature stability of μ_Δ, and low eddy-current losses. Because of this, they have rapidly ousted expensive strip- and ribbon-wound cores from cold-rolled silicon steel and Permalloy.

Typically, ferrite cores have $B_{sat} = 1\,500$ to $3\,000$ gauss, $B_{r0} = 80$ to $1\,000$ gauss, and $\Delta B_{max} = 700$ to $2\,000$ gauss. As a way of reducing the remanence, they are made with an air gap with a width equal to from 0.001 to 0.0001 of the mean core length and placed in an auxiliary demagnetizing field H_0 of about 1 to 2 gauss [31]. In such cases, it is preferable to employ square-loop ferrites for which $B_{r0}/B_{sat} > 0.8$.

To take care of magnetic hysteresis in designing a pulse transformer, the absolute permeability μ_a in Eq. (2.42) should be replaced with μ_Δ, and the maximum increment in magnetic induction, ΔB, should be limited by increasing w_1 or S_c.

Eddy currents. Fast variations in the magnetic flux threading the core of a pulse transformer give rise to eddy currents which tend

* Ferrite is a powdered, compressed and sintered magnetic material consisting chiefly of ferric oxide combined with one or more other metals. Examples of ferrite compositions include nickel ferrite, nickel-cobalt ferrite, manganese-magnesium ferrite and some others.—Translator's note.

to weaken the main magnetic field. To make up for this effect, an additional magnetizing field must be applied. In the equivalent circuit of Fig. 2.20*b*, the increase in the resultant magnetic flux is represented by the increase in the magnetizing current i_μ produced by additional parallel *RL*-networks [4]. The magnetizing inductance can be replaced by an apparent inductance L_{ap} which is defined by Eq. (2.42), with μ_a replaced by the apparent permeability μ_{ap} which is a function of the pulse duration. The resistivity of ferrite cores is about 10^7 times as high as that of cores from cold-rolled steel and Permalloy. Because of this, eddy currents in ferrite cores are negligibly small, and there is practically no reduction in the value of μ_Δ.

The pulse permeability of ferrite cores is reduced by what is known as magnetic viscosity which is responsible for the time delay with which the sense of magnetization in a ferrite core can be reversed. According to data sheets which quote the values of μ_p with allowance for magnetic viscosity, the reduction in pulse duration from 3 to 5 or 1 μs, the permeability goes down by no more than 10%.

Leakage flux and displacement currents in the windings. In a real transformer, an electric field is set up between the windings and also between each winding and the core (frame). When a pulse voltage or current is impressed on the transformer, the intensity of this field varies at a very high rate, giving rise to displacement currents which shift the balance of currents in the transformer. The magnitude of displacement currents is decided by the distributed capacitance C_0 of the transformer windings. The value of C_0 can be found by treating the winding layers as the parallel plates of a capacitor [3].

In the equivalent circuit of a pulse transformer, the distributed capacitances must be replaced by lumped equivalent (dynamic) capacitances whose value will depend on the point of connection in the circuit, turns number, and the arrangement of the windings. For practical purposes, it is reasonably accurate to place one equivalent (dynamic) capacitance in parallel with each transformer winding. The dynamic capacitances can be calculated by assuming that the electrostatic energy stored in the windings is equal to that stored by the equivalent lumped capacitor, and also that the potentials vary linearly along the length of the windings [5]. For a real pulse transformer, the equivalent capacitances can be found by experiment in the same manner as the equivalent distributed capacitances of inductors.

Leakage flux in a pulse transformer is that part of the total magnetic flux which is diverted from the main path and links with various groups of turns adjacent to a given turn. Because of this, the voltage transformation ratio is reduced in comparison with the value given by Eq. (2.42) and becomes dependent on the signal spectrum and the character of load.

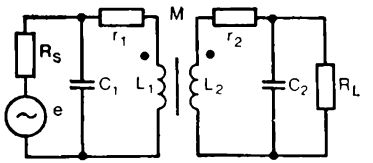


Fig. 2.22

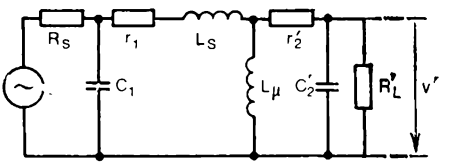


Fig. 2.23

If we assume, as a rough approximation, that the current is the same in all the turns of a given winding, the magnetic fluxes threading the transformer windings will differ as much as the coupling factor k_c differs from unity:

$$k_c = M/\sqrt{L_1 L_2} \tag{2.45}$$

where M is the coefficient of mutual inductance between the windings.

Figure 2.22 shows an equivalent circuit of a pulse transformer with allowance for displacement currents and leakage flux. It also shows the series loss resistances r_1 and r_2 of the windings and the internal resistance R_s of the pulse source. The capacitors C_1 and C_2 include the capacitive components, C_s and C_L, of the pulse generator and load, and the dynamic capacitances C_d of the transformer windings:

$$C_1 = C_s + C_{d1},\ C_2 = C_L + C_{d2} \tag{2.46}$$

The principal components of the dynamic capacitances are the equivalent capacitance between the primary and secondary windings, $C_{d1,2}$ and the equivalent capacitances, $C_{d1,0}$ and $C_{d2,0}$, between the windings and the transformer core (frame) usually held at zero (ground) potential:

$$C_{d1} = C_{d1,0} + C_{d1,2},\quad C_{d2} = C_{d2,0} \tag{2.47}$$

Equivalent circuit of a real pulse transformer. The relations describing the equivalent-circuit parameters depend on the point of connection of the series inductance characterizing magnetic leakage. Since, however, the leakage inductance L_{ik} in a pulse transformer is small, the difference in parameters between a real transformer and its equivalent circuit is negligible. Therefore, the leakage inductance $L_{lk} = L_s$ is most often placed ahead of the magnetizing inductance L_μ (Fig. 2.23) given by Eq. (2.42) for $\mu = \mu_p$. The parameters of this equivalent circuit are [6]:

$$\begin{aligned} &L_{lk} = (1-k_c^2)\,L_1 \approx 2L_1(1-k_c),\quad L_\mu = k_c^2 L_1 \approx L_1 \\ &C_2' = (n^2/k_c^2)\,C_2 \approx C_2 n^2 \\ &r_2' = r_2 k_c^2/n^2 \approx r_2/n^2,\quad R_L' = R_L k_c^2/n^2 \approx R_L/n^2 \end{aligned} \tag{2.48}$$

The approximate relations defined by Eqs. (2.48) are developed on the assumption that in properly designed pulse transformers the

coupling coefficient will usually be greater than 0.99. Therefore, the leakage inductance does not normally exceed $0.01L_\mu$ and can range from a few tenths of a microhenry to tens of microhenrys, according to type of transformer.

The total winding resistance $r_1 + r_2' \approx 2r_1$ will usually be a few per cent of the sum $R_L' + R_s'$, and it is ordinarily neglected. The effects of hysteresis and eddy currents are taken care of by placing the magnetizing inductance L_μ in the parallel arm of the circuit (see Fig. 2.23).

Transients in the above circuit are described by fourth-order differential equations. For this reason, fast processes (such as may be associated with the leading and trailing edges of pulses) are analyzed separately from slow processes (associated with the pulse top) in approximate sketch calculations. Where high accuracy is important, the equations describing the state of a circuit are developed and solved by a computer [28].

Tilt of the pulse top. For an approximate evaluation of the tilt of a pulse of duration τ, the equivalent circuit of Fig. 2.23 is re-arranged to drop the stray capacitances C_1 and C_2' and the leakage inductance L_{1k}, because their effect on the shaping of the pulse top is negligible in comparison with the magnetizing inductance L_μ. The equivalent circuit may further be simplified by dropping the series loss resistances of the windings which, as already noted, are negligibly small. With these simplifications, the pulse tilt is stated in fractional form as

$$\delta V = \tau/\tau_\mu \tag{2.49}$$

where

$$\tau_\mu = L_\mu/R_{eq}$$

and

$$R_{eq} = R_s \parallel R_L'$$

Distortion of the leading and trailing edges. For an approximate evaluation of distortion suffered by the leading and trailing edges of the output pulse, the parallel arm containing the magnetizing inductance L_μ may be dropped from the circuit of Fig. 2.23 because variations of magnetizing current over these relatively short intervals may be ignored. Because L_s is small, C_1 and C_2' are customarily replaced by a single capacitance, $C = C_1 + C_2'$, which is placed instead of the larger of the two. For step-up transformers ($n > 1$), C_1 is usually less than C_2', whereas for step-down transformers ($n < 1$), $C_1 > C_2'$. In practice, it often happens that after the input pulse ceases, the source impedance R_s and the load impedance R_L are sharply changed, as this occurs in blocking oscillators (see Chap. 8). Therefore, the parameters R_s, C and R_L' in the equivalent circuits

for the leading and trailing edges are different in the general case.

The pulse rise time is a minimum for an underdamped (oscillatory) response, but damped oscillations (ringing) are superimposed on the pulse top then.

The resultant waveshape of the output pulse can be determined from analysis of simplified equivalent circuits by recourse to curve-fitting for the slow and fast transients. As is shown in [31], this involves multiplying together the normalized transient responses for the leading and trailing edges of the pulse.

In designing a pulse transformer, it is usually sought to minimize both L_{1k} and C by reducing the number of turns in the windings ($w_{1,2} \leqslant 100$ turns), setting the transformation ratio close to unity, using single-layer windings, and placing them on a common limb. However, the behaviour of a pulse transformer is such that a decrease in L_{1k} (or C) will usually bring about a nearly inevitable increase in C (or L_{1k}). The distributed component of C is inversely proportional to the thickness of the insulation between the windings and between each winding and the core, whereas the leakage inductance is directly proportional to the insulation thickness. Also, L_{1k} is proportional to the square of the number of turns and the mean turn length.

In some applications, as in, say, blocking oscillators where the load is nonlinear, an important factor is the waveshape of the output voltage following the output pulse, because it is affected by the decrease in the magnetizing current. As already noted, if the shunt capacitances were non-existent, the pulse tail would decay exponentially. If, however, R'_L and R_s are high, the presence of the shunting capacitance C will result in damped oscillations (ringing). To avoid them, it might be necessary to include an additional shunt resistor.

2.7. DELAY LINES FOR PULSE SIGNALS

The term *delay line* refers to a passive two-port which delivers an output signal with a certain time delay relative to the input signal. An example of the passive two-ports displaying this property is a nondispersive (distortion-free) distributed-constant transmission line.

The group velocity u_g and the phase velocity u_p of a signal propagating down a nondispersive transmission line are given by

$$u_g = u_p = u = 1/\sqrt{L_0 C_0} \tag{2.50a}$$

where L_0 and C_0 are the inductance and capacitance per unit length of the line. Hence, the delay T_0 of a signal per unit length of the line is

$$T_0 = u^{-1} = \sqrt{L_0 C_0} \tag{2.50b}$$

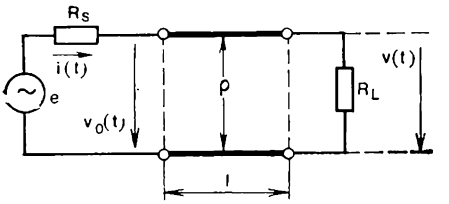


Fig. 2.24

Real transmission lines are, as a rule, dispersive, which implies that the individual spectral components of a pulse signal propagating in a real line differ in phase velocity. The net result is distortion in the waveshape of the delayed signal. Also, the output signal is reduced in amplitude because of losses in the line.

Apart from the time delay, the basic characteristics of an electric delay line are the form of the transient response (as it determines the distortion in the signal), attenuation, characteristic (or wave) impedance, physical dimensions, maximum permissible voltage on the elements, stability of delay with temperature and time, accuracy of delay setting, and cost.

In the pulse field, delay lines are used on a very wide scale. This includes coders and decoders, pulse selectors, pulse-width and pulse-rate stabilizing networks, and various digital circuits. Using a step input and a delay line, rectangular pulses of regular shape can readily be generated on relatively small-value load impedances arranged into simple and reliable circuits.

Delay line as a circuit element. If we place a homogeneous and nondispersive delay line of length l and characteristic impedance $W = \rho = \sqrt{L_0/C_0}$ between a generator supplying a voltage $e(t)$ and having an internal impedance R_s, and a load R_L (Fig. 2.24) then at $R_s \neq \rho$ and $R_L \neq \rho$ the voltage at the line input, $v_0(t)$, will be the sum of the voltages contributed by forward and backward waves set up in the line after a time τ equal to twice the delay T in one direction:

$$\tau = 2T = 2T_0 l \tag{2.51}$$

If the duration of the input pulse is $\tau_{in} < \tau$, then a train of pulses $e(t)$ with a damped amplitude will be produced at the input (and output) of the delay line. If, on the other hand, $\tau_{in} > \tau$, then $v_0(t)$ will be a stretched straircase waveform.

When the signal source is perfectly matched to the line ($R_s = \rho$), as is usually the case in practice, the Laplace transform of $v_0(t)$ will be written [4]:

$$v_0(s) = 0.5\,[1 + K_{back} \exp(-2sT)]\, e(s) \tag{2.52}$$

where

$$K_{back} = (R_L - \rho)/(R_L + \rho)$$

In the circumstances, no multiple reflections occur in the line, and $v_0(t)$ contains two components, the second (complementary) component appearing at time $t = 2T$.

From a practical point of view, the following cases are of interest.

Transmission of pulses with a time delay. It is arranged that the load is matched to the characteristic impedance of the line ($R_L = \rho$), so there is no reflected or backward wave ($K_{back} = 0$), and the original time function of the transform given by Eq. (2.52) has the form

$$v_0(t) = 0.5e(t) \tag{2.53a}$$

In transmission of pulses with a time delay, the operating voltage is v_{out} which has the same waveshape as, but is shifted by a time $\tau = T$ from, the input voltage:

$$v_{out}(t) = 0.5e(t - T) \tag{2.53b}$$

Sometimes, the impedances are matched in a different way, so that $\rho = R_L \gg R_s$. Then,

$$v_{out}(t) = e(t - T)\, R_L/(R_L + R_s) \approx e(t - T) \tag{2.53c}$$

which is twice as great as the signal defined by Eq. (2.53b). It is pointed in [31], however, that if the delay line is not precisely matched to the load in this case, the waveform of the output pulse will be distorted more.

Pulse-shaping by a short-circuited delay line. Now $R_L = 0$, so the voltage of the wave reflected from the load end of the line is equal in magnitude and opposite in sign to that of the forward wave, and the input voltage is proportional to the difference between step input functions shifted by $2T$ from each other:

$$v_0(t) = 0.5\,[e(t) - e(t - 2T)] \tag{2.54}$$

Pulse-shaping by an open-circuited delay line. In this case, $R_L = \infty$, $K_{back} = 1$, and

$$v_0(t) = 0.5\,[e(t) + e(t - 2T)] \tag{2.55}$$

The last two cases are used in practice especially often. Consider them in more detail.

The waveforms of the voltage $v_0(t)$ and current $i(t)$ delivered to the line by a pulse generator driven by a rectangular input $e(t) = E$ in the last two cases are shown in Fig. 2.25. Each time a drive pulse of a relatively long duration ($\tau_{in} > 2T$) is applied to the pulse generator, the output waveform is one of two bipolar pulses of duration $\tau = 2T$ in either case. The difference is that with a short-circuited delay line the bipolar waveform is the voltage pulses $v_0(t)$

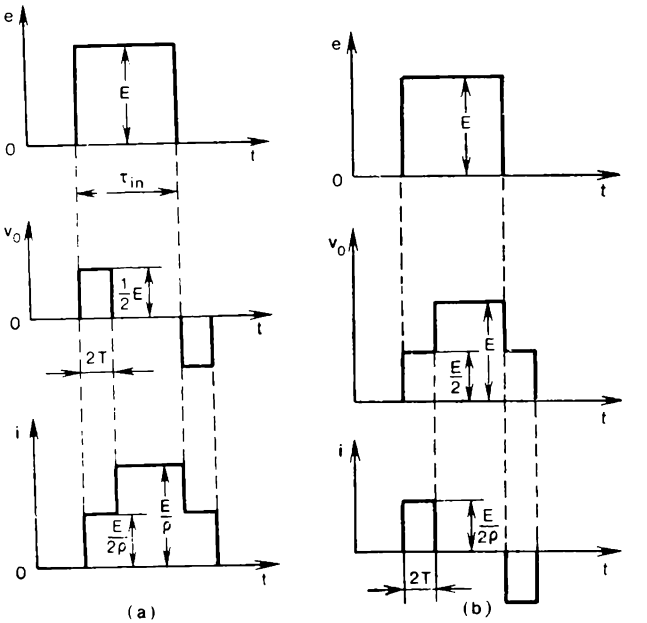


Fig. 2.25

at the line input (see Fig. 2.25*a*), whereas with an open-circuited delay line the bipolar waveform is the current pulses $i\,(t)$ in the circuit from the pulse generator to the delay line (see Fig. 2.25*b*). Therefore, the load should be placed across the input terminals of the line in the former case, and between the pulse generator and the line in the latter (Fig. 2.26). Accordingly, there will be a difference in conditions for impedance match at the line input required to avoid multiple reflections:

$$R_{eq} = \begin{cases} R_s \,||\, R_L = \rho \text{ for Fig. } 2.26a \\ R_s = R_L = \rho \text{ for Fig. } 2.26b \end{cases} \tag{2.56}$$

Thus, the applicability of a short- or an open-circuited delay line to pulse-shaping depends above all on whether or not the conditions for

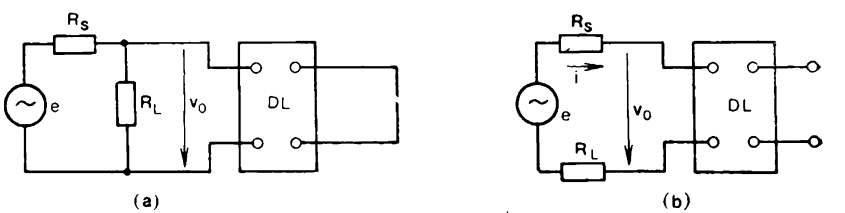


Fig. 2.26

impedance match at the input (operating) end of the line, defined by Eq. (2.56), can be satisfied.

To sum up, a pulse generator should contain a short- or an open-circuited delay line driven via a switch by a step or rectangular voltage with a step duration longer than twice the delay time of the line. In its effect, the delay line will be analogous to a pulse-shortening RC- or RL-networks, but pulses with a duration $\tau = 2T$ will be closer to a rectangular waveshape.

Accuracy of impedance match at the line input. The accuracy with which the characteristic impedance of the line must be matched to the equivalent load impedance R_{eq} can be evaluated from the maximum permissible amplitude of voltage or current over the interval $2T < t < 4T$. If the drive function is a step voltage, $e = E$, and $K_{back} = \mp 1$ (for a short- and an open-circuited line, respectively), then, given an imperfect match ($R_s \neq \rho$), we find [4] that the amplitude ratio for the first two voltage pulses, δv_{sc}, at the input of a short-circuited line and for the first two current pulses, δ_{ioc}, in an open-circuited line will be equal:

$$\delta v_{sc} = -\delta i_{oc} = -K_f \tag{2.57}$$

where

$$K_f = (R_s - \rho)/(R_s + \rho)$$

Thus, if the amplitude of a succeeding signal should not exceed 10% of the amplitude of the first pulse, the mismatch should lie within the limits $0.82 \geqslant R_{eq}/\rho \geqslant 1.22$.

Effect of reactive load. If the load is not purely resistive, the condition for impedance match at the input end of a homogeneous line cannot be satisfied, and there will be, as a consequence, a distortion in the waveshape of the output pulses. It can be shown [9] that in cases where the load is a resistance shunted by an inductance or placed in series with a capacitance, rectangular pulses can be produced by lines whose characteristic impedance varies hyperbolically or parabolically. If the load resistance is shunted by a stray capacitance, the leading and trailing edges of the output pulse can be corrected by including a network which generates a voltage spike restoring the pulse shaped by the delay line to a rectangular waveform [9].

Coaxial-cable delay line. Shielded r.f. cables are often used as precision pulse-shaping and delaying devices. But the maximum delay time of such lines is limited to the nanosecond range, because the phase velocity of a signal propagating in an r.f. cable is appreciable.

When the centre and outer conductors of a coaxial line are separated by a homogeneous dielectric with dielectric constant ε, the low-frequency phase velocity u_p is

$$u_p = c_0/\sqrt{\varepsilon} = 1/\sqrt{L_0 C_0} \tag{2.58}$$

where c_0 is the velocity of light in a vacuum. Since $c_0 = 3 \times 10^8$ m s^{-1} the low-frequency delay time per unit length of such a cable is

$$T_0 = 0.0033 \sqrt{\varepsilon} \qquad (2.59)$$

For the most frequently used low-loss dielectrics (polystyrene, polyethylene, or PTFE), $\varepsilon = 2.3$ and $T_0 = 0.005$ μs m^{-1}. To obtain a delay of 1 μs, it would be necessary to use a cable 200 m long, which is impracticable. Equations (2.58) and (2.59) hold at frequencies where the dispersive behaviour of the line (variations in the phase velocity with frequency) may be neglected.

In the nanosecond range, the dispersive behaviour of coaxial lines is responsible for an increase in the pulse rise and fall times with the increase in cable length [31].

For conventional lines, the characteristic impedance is very low ($\rho = 50$ to 125 Ω), so they are inconvenient to match to load impedance.

Spiral delay lines. A very simple way to extend the delay time per unit length of a line would be to use a dielectric material having a dielectric constant of as high as 150 (such as ceramic materials ordinarily used in capacitors). This would, however, reduce the characteristic impedance of the line to a few ohms, and it would be extremely difficult to match the delay line to the load and pulse generator. A better way out is to use a transmission line with a helical inner conductor—it serves to increase both the inductance per unit length and the characteristic impedance of the line. The delay time per unit length of a spiral delay line increases with increasing diameter of the helix and decreasing thickness of the insulation between the helix and the outer conductor made in the form of braid from insulated wires connected together only at the ends in order to minimize eddy-current losses.

For the most-commonly used spiral delay lines with a dielectric core, the time delay per unit length, T_0, ranges from 0.18 to 1.8 μs m^{-1}, and the characteristic impedance is $\rho = 1$ kΩ. When the inner conductor is a ferrite rod ($\mu = 4$), T_0 and ρ are increased two-fold.

Among the drawbacks of spiral delay lines is the decrease in the characteristic impedance and delay time with increasing frequency. This is attributed to the decrease in inductance per unit length because of the increase in phase shift between current pulses from turn to turn in the helix. Also, the attenuation introduced by the line increases with rising frequency. These factors lead to distortion in the pulse waveshape, mainly as an extended pulse rise time. This is the reason why spiral delay lines are used where the delay time need not exceed several microseconds.

Strip lines. In integrated-circuit (IC) technology, nanosecond delays are obtained with delay lines built from strip-type transmission lines of which there are two general forms, shielded (or balanced)

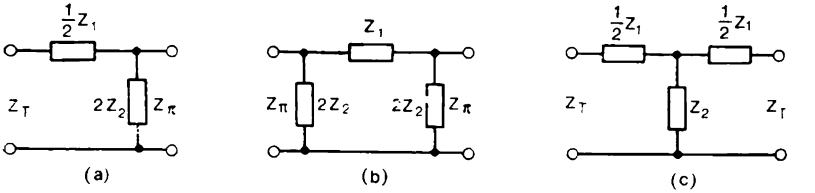


Fig. 2.27

and unshielded (or unbalanced), the latter more frequently called *microstrip lines.* A shielded strip line (or simply, a *strip line*) consists of a strip conductor between two ground planes in the form of metal plates separated by a dielectric material. A microstrip line is a strip (or flat) conductor above a single ground plane. As with coaxial cables, the characteristic impedance of strip-type transmission lines depends on their mechanical dimensions and effective permeability.

Pulse-shaping by lumped-constant networks. In cases where the delay or pulse duration exceeds several microseconds, distributed-constant delay lines become unsuitable because of their mechanical dimensions and the attenuation they introduce. Instead, recourse is had to a chain of lumped-constant LC-sections.

Ladder networks. Most frequently, a lumped-constant delay line is obtained by dividing a distributed-constant delay line into sufficiently short lengths and replacing each length by a section whose L and C parameters are equal to the product of the respective parameters per unit length, averaged over the section. Pulse-shaping circuits thus obtained have come to be known as *ladder networks.* For a purely resistive load, use is made of homogeneous ladder networks in which the main sections are identical. Generally, a ladder network may be a chain of L-sections (Fig. 2.27*a*), pi-sections (Fig. 2.27*b*), and/or T-sections (Fig. 2.27*c*).

If the series element is an inductor ($Z_1 = j\omega L$) and the shunt (parallel) element is a capacitor ($Z_2 = 1/j\omega C$), the product Z_1Z_2 will be frequency-independent (Fig. 2.28), that is,

$$k = Z_1Z_2 = L/C = \rho^2 \tag{2.60}$$

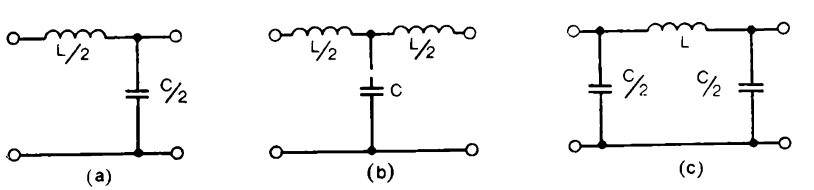


Fig. 2.28

Therefore, the sections shown in Fig. 2.28 are called constant-k sections.

For a tandem connection of n constant-k sections loaded into an ohmic load, the characteristic impedances can only be matched between the inner sections, whereas the input section connected to the generator and the output section connected to the load will remain unmatched. Because of this, the signal passing through a ladder network will suffer amplitude-frequency distortion in the outer sections in addition to the distortion caused by the nonlinearity of the resultant phase response.

Consider a single constant-k section loaded into a resistive load $R_L = \rho$. From analysis of its transient response we can see that the delay T_1, defined as the time interval required for the output voltage to rise to half its steady-state value, and the pulse rise time τ_{r1}, defined as the time interval required for the pulse to rise from 10% to 90% of its steady-state amplitude, are given by

$$\begin{aligned} T_1 &= 1.07\sqrt{LC} \\ \tau_{r1} &= 1.13\sqrt{LC} \end{aligned} \tag{2.61}$$

In the circumstances, the overshoot amounts to 8% of the steady-state amplitude.

Experiments [5] have shown that for n matched constant-k sections placed between the generator and load we may write

$$\begin{aligned} T_n &= nT_1 \\ \tau_r &= \tau_{r1} n^{1/3} \end{aligned} \tag{2.62}$$

Combining Eqs. (2.61) and (2.62) yields the condition for selecting the number n of constant-k sections, with the pulse duration τ_x and the pulse rise time τ_r given in advance:

$$\begin{aligned} n &= 0.39\,(\tau_p/\tau_r)^{3/2} \\ n &= 1.1\,(T_n/\tau_r)^{3/2} \end{aligned} \tag{2.63}$$

To sum up, in order to shape a pulse with a relative rise time of 10%, we would need a line containing at least eleven constant-k sections.

To reduce the number of sections and to improve the pulse waveshape, the constant-k section must be transformed so as to increase the delay provided by a single section without an appreciable increase in the pulse rise and fall times, to improve the linearity of the phase response, and to decrease the frequency dependence of the characteristic impedances. These goals are achieved by what are called *m-derived sections* (Fig. 2.29). Their parameters are chosen such that the characteristic impedance and the cut-off frequency are equal to the respective quantities of the prototype constant-k section.

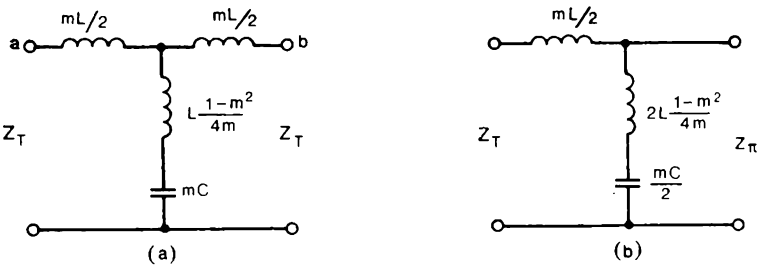


Fig. 2.29

The value of m determines the behaviour of the network. With $m = 1$, m-derived sections reduce to constant-k sections. For the inner sections of a multisection line, the optimal value of m is 1.27, because the phase response is then maximally linear. For the outer (terminating) sections, the optimal value is chosen to be $m = 0.6$ (Fig. 2.30*a*) so as to obtain maximally constant values for the input impedances.

At $m > 1$, the inductance in the parallel (shunt) arm of an m-derived section becomes negative and can be implemented by introducing magnetic coupling between the series inductors of the T-section (Fig. 2.30*b*). At $m = 1.27$, $k_c = 0.23$, $L_m = 1.02L$, and $C_m = 1.27C$. For an inner m-derived section with $m = 1.27$:

$$T_1 = 1.2\sqrt{LC}$$
$$\tau_{r1} = 1.15\sqrt{LC} \tag{2.64}$$

The total delay provided by a line consisting of a number n of m-derived sections and its transient response time are given by Eq. (2.62), as for constant-k sections. Combining Eqs. (2.62) and (2.64), we obtain

$$n = 0.94\,(T_n/\tau_r)^{3/2}$$

or (2.65)

$$n = 0.33\,(\tau_p/\tau_r)^{3/2}$$

From a comparison of Eqs. (2.63) and (2.65) it follows that, given the same ratio T_n/τ_r or τ_p/τ_r, the number of m-derived sections will

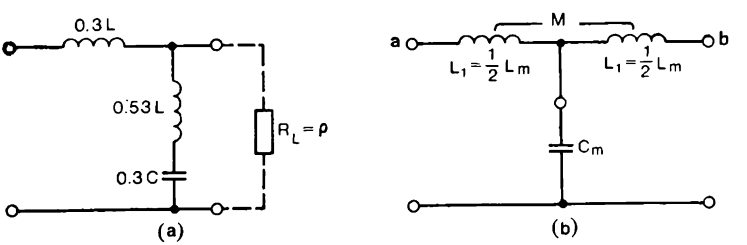


Fig. 2.30

be by 16% smaller than the number of constant-k sections. For $\tau_r/\tau_p = 0.1$, nine instead of eleven sections will suffice. Also, the overshoot with an m-derived section is about half as great as with a constant-k section.

Pulse-shaping one-ports containing resonant circuits. When ladder-network delay lines are used to shape rectangular pulses, they suffer from a number of drawbacks, namely: many sections have to be used in order to obtain sufficiently fast rise and fall times, and there is a large amount of ringing superimposed on the pulse top. This is the reason why in applications where the amount of ringing and the size of a pulse-shaping circuit are critical, resort is made to a pulse-shaping network consisting of several resonant circuits which differ in resonant frequency and characteristic impedance. The nominal resonant-circuit capacitance and inductance can be evaluated by expanding the Laplace transforms of the input admittance or impedance into partial fractions or by approximating the transfer functions by trigonometric sums or polynomials.

The pulse waveshape might be improved and the number of sections reduced by assigning different nominal values to one or both elements (L and C) of each resonant circuit even in a homogeneous pulse-shaping delay line. This would however complicate their manufacture, especially in the case of an increase in the power of output pulses in comparison with ladder-network delay lines (for a more detailed discussion, refer to [4] and [31]).

Ultrasonic delay lines. Ultrasonic delay lines give relatively long delays (from 10 μs to several milliseconds). For its operation, an ultrasonic delay line depends on the propagation of mechanical vibrations at ultrasonic frequencies through a solid (fused quartz, ceramics, magnesium alloys and the like). The vibrations are induced and picked up by directional piezoelectric transducers made in the form of thin quartz plates and installed at right angles to the propagation of sonic waves. As a way of reducing the size of ultrasonic delay lines, use is made of sonic waveguides with multiple reflections from the side faces.

The frequency characteristic of ultrasonic delay lines is matched to the spectrum of the delayed pulses by firstly converting video pulses to r.f. pulses in an r.f. generator modulated by the video pulses to be delayed. The electric signals picked up by the receiving transducer are amplified and detected [31].

REVIEW QUESTIONS AND PROBLEMS

2.1. Prove the equality of the volt-second areas of voltage at the output of an RC-network (Fig. 2.1) driven by a single voltage pulse.

2.2. Determine distortion in single rectangular and triangular voltage pulses on passing through a d.c. blocking RC-network (Fig. 2.1).

2.3. Determine the waveform of the voltage at the output of a d.c. blocking RC-network (Fig. 2.1) driven by a train of rectangular pulses.

2.4. Determine the waveform of the voltage at the output of a pulse-shortening RC-network driven by rectangular pulses with infinitesimal and finite (in the case of linear approximation) rise and fall times.

2.5. Consider the effect of the stray parameters in a pulse-shortening RC-network (Fig. 2.8) driven by rectangular and triangular pulses on assuming: (a) $R_s \neq 0$, $C_s = 0$; (b) $R_s = 0$, $C_s \neq 0$; (c) $R_s \neq 0$, $C_s \neq 0$.

2.6. Show that, with the given amplitude E of input rectangular pulses, time constant R_sC_s and pulse duration τ_p, the network of Fig. 2.8 has at its output optimal parameters $R = R_{opt}$ and $C = C_{opt}$ such that the output pulses have a maximum amplitude.

2.7. Consider the voltage waveform at the output of an RC-network loaded by a capacitor (Fig. 2.10), driven by rectangular and triangular pulses with $\tau \ll T$ and $\tau \gg T$.

2.8. Consider the effect of the load impedance R_L (Fig. 2.12) on the output voltage waveform in a network driven by rectangular and triangular pulses with $\tau \ll T$.

2.9. Explain why the accuracy of differentiation and integration performed by simple RC-networks cannot be improved.

2.10. Explain how the use of a compensating voltage can improve the accuracy of differentiation and integration.

2.11. Draw a block diagram of an idealized feedback operational amplifier, derive an analytical expression for the transfer function $K(s)$, and show when an operational amplifier displays high efficiency.

2.12. Determine the equivalent input impedance of an idealized operational amplifier; explain when an apparent short-circuit exists at the input of an operational amplifier, such that $v_{in} = 0$ and $i_{in} = 0$.

2.13. Define the transfer function for an operational amplifier at $R_{in} \neq \infty$.

2.14. Determine the equivalent output impedance, $R_{out,eq}$, of an operational amplifier for $R_{out} \neq 0$, using the general method for evaluating the differential impedance between two terminals of a linear multi-port, based on the relation $R_{out,eq} = e/i$, where e is the emf of the source connected to these terminals, and i is the consumption current.

2.15. Draw a block diagram for a differentiating opamp; form the transfer function with the drive applied to the direct and inverting inputs; demonstrate the effect of $R_{in} \neq \infty$.

2.16. Examine the waveform of v_- in the network of Fig. 2.19 driven by a step input function for (a) the overdamped case and (b) the critically damped case.

2.17. Define the manner in which the voltage varies across a shock-excited resonant circuit (Fig. 2.19), assuming an oscillatory response.

2.18. Develop an equivalent circuit for an idealized pulse transformer ($r_1 = r_2 = 0$, $L_{lk} = 0$, $C_s = 0$, μ = constant, $d = 0$, $L_1 \neq \infty$); evaluate the effect of L_1 on the step-function response.

2.19. Evaluate the effect of magnetic hysteresis on the performance of a pulse transformer.

2.20. Evaluate the effect of eddy currents on the performance of a pulse transformer.

2.21. Evaluate the effect of leakage flux and displacement currents on the performance of a pulse transformer.

2.22. Form an equivalent circuit for a pulse transformer with allowance for the effect of magnetic hysteresis, eddy currents, leakage flux, and displacement currents.

2.23. Evaluate the tilt of a rectangular pulse passed by a pulse transformer.

2.24. Evaluate distortion in the leading and trailing edges of a rectangular pulse passing through a pulse transformer.

2.25. Derive Eq. (2.52).

2.26. Derive Eqs. (2.53) through (2.55).

2.27. Develop networks to shape pulses in accordance with Eqs. (2.54) and (2.55).

2.28. Determine the limit of mismatch between the characteristic impedance of a delay line and an equivalent load impedance.

2.29. List the basic types of constant-k and m-derived sections of pelay lines and demonstrate their basic differences in shaping and dassing input pulses.

3 SWITCHES

3.1. GENERAL

A switch is a device that has two operating states: ON and OFF, differing in the value of an output parameter which may be r_{out}, v_{out}, i_{out}, etc. A change of state occurs in response to a signal (or signals) applied to one or several inputs.

An elementary switch (a *switching element*) is a nonlinear device the two operating states of which differ in output resistance. Switching elements serve to close and open electric circuits and are the basic components of any switch. In addition to a switching element, a switch may contain power supplies, signal amplifiers, and passive elements such as resistors, capacitors, etc.

An ideal switching element can be simulated by a lag-free mechanically operated switch (Fig. 3.1) which has an infinite resistance in the unoperated (open or off) state, and an infinitesimal resistance in the operated (closed or on) state. Whatever the parameters of the switched circuit, R_{eq} and e_{eq}, the following relations hold:

$$V_c = 0,\ I_o = 0,\ \tau_{ON} = \tau_{OFF} = 0 \qquad (3.1)$$

where V_c is the voltage across the closed switching element, I_o is the current through the open switching element, τ_{ON} and τ_{OFF} are the times required for the switch to change state.

Pulse circuits use static switches which are nonlinear electron devices, such as crystal diodes, transistors, vacuum or gas-filled tubes, thyristors, etc. In the switching mode, such devices have two states substantially differing in output resistance. In the ON state, this output resistance is a small fraction of the equivalent resistance of the switched circuit, whereas in the OFF state it is many times the equivalent resistance. Therefore, the conditions defined by Eqs. (3.1) are satisfied only approximately.

Now we shall take up the basic types of crystal diodes and bipolar transistors employed as switches and the main circuit configurations devised for the purpose. In this chapter, we shall limit ourselves to single-input circuits, and multi-input circuits will be discussed in Chap. 5.

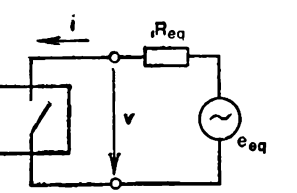


Fig. 3.1

3.2. DIODE SWITCHES

Pulse diode types. Both point-contact and junction types of diode are used as switches in pulse circuits, the junction diodes being of either the alloyed or the diffused type. Point-contact diodes have a small transition area and a correspondingly low barrier capacitance (C_b = 1 or 2 pF). In comparison with junction diodes, however, they have a lower stability and reliability, a limited power dissipation, and a lower peak inverse voltage rating. Alloyed junction diodes show a better performance, but they have an increased barrier capacitance (C_b = 10 to 20 pF) owing to a larger transition area. Diffused junction diodes combine a good measure of the advantages offered by both point-contact and alloyed junction diodes and are therefore more promising, especially planar structures [30]. In addition to diffused diodes, integrated circuits use bipolar transistor structures connected as diodes [24] and *Shottky-barrier diodes* based on a metal-semiconductor junction*. In a Shottky diode [42], the current is constituted by the flow of majority carriers, and no minority carriers are accumulated in the base region. Because of this, the time lag (inertia) associated with Shottky diodes is solely due to the time required to recharge the small barrier capacitance, C_b (from several tenths to a few picofarads).

Another variety of pulse diodes having an increased storage time are charge-storage diodes.

Static Models of Crystal Diodes

The static V/I characteristic of a crystal diode is shown in Fig. 3.2. The switching diode is in the ON state in the low-resistance part of the forward V/I characteristic (V_f, I_f), and in the OFF state in the high-resistance part of the reverse V/I curve ($-I_r$, $-V_r$). As the cutoff voltage is increased, the diode moves into the breakdown region ($-V_{bd}$, $-I_{bd}$) where the dynamic resistance is small. Most pulse diodes are not allowed to operate in this region. An exception is silicon Zener diodes for which the breakdown region is an operating region where a Zener diode can

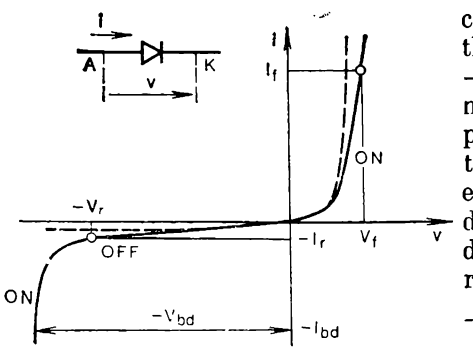


Fig. 3.2

* Also known as *surface-barrier* or *hot-carrier* diodes.—Translator's note.

double as a switching element connected in series with a d.c. voltage source.

The models used to represent the static behaviour of a crystal diode vary in complexity according to the degree of detail sought.

The simplest model is an *ideal diode* (Fig. 3.3*a*) which approximates the *V/I* characteristic of a real diode by straight-line segments, one coinciding with the negative *v*-axis and the other with the positive *i*-axis (Fig. 3.3*b*). The ideal diode D_i may be looked upon as a mechanical switch (see Fig. 3.1) whose contacts are closed when $i > 0$, and open when $v < 0$. At the instant when the contacts close, $v(t) = 0$, and at the instant when they open, $i(t) = 0$.

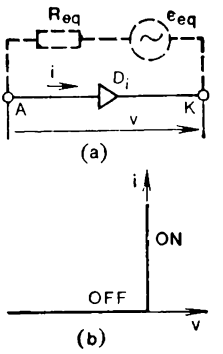


Fig. 3.3

Piecewise-linear models. In contrast to the ideal diode, the piecewise-linear model allows for the finite resistances in the ON and OFF states. An approximating *V/I* characteristic consisting of three straight-line segments is shown in Fig. 3.4. Segment *1* approximates the low-resistance portion of the forward characteristic, segment *2* the high-resistance part of the reverse characteristic, and segment *3* the breakdown region. For each segment there is a linear equivalent circuit (Fig. 3.5); by inspection of each circuit the respective linear portions of the *V/I* characteristic can be written:

$$v = e_0 + r_f i \quad \text{for} \quad i > I_{c\text{-}o}$$
$$i = -I_0 + v/r_r \quad \text{for} \quad -V_{bd} < v < V_{c\text{-}i}$$
$$v = e_{bd} + r_{bd} i \quad \text{for} \quad i < -I_{bd}$$

where e_0, I_0, and e_{bd} are the initial values of the forward voltage, reverse current and breakdown voltage, respectively, defined as the coordinates of the points where the approximating straight-line segments (or their extensions) intersect the coordinate axes; r_f, r_r, r_{bd} are the equivalent forward, reverse and breakdown resistances close to the respective dynamic resistances averaged over the respective portions; $(V_{c\text{-}o}, I_{c\text{-}o})$ and $(-V_{bd}, -I_{bd})$ are the coordinates of the breakpoints *A* and *B* (Fig. 3.4) defining the conditions for a diode switch to change state.

Figure 3.6*a* shows two likely locations for the breakpoint *A*. In the first case, the breakpoint lies on the forward characteristic of the diode, with its position A_1 chosen such that in the ON state (to the right of A_1) the diode resistance is smaller than the equivalent resistance R_{eq} of the switched circuit, whereas in the OFF state (to the left of A_1) it is greater than R_{eq}; in other words, at point A_1 the dynamic resistance of the diode is $r_{bp} \approx R_{eq}$. Since in the switching mode it is required that $r_f \ll R_{eq} \ll r_r$, the value of r_{bp} ought

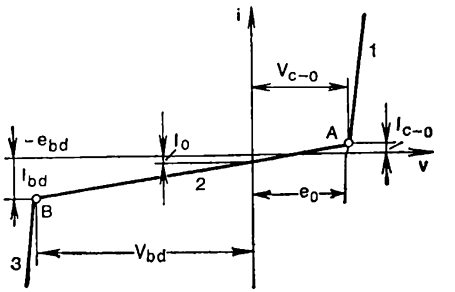


Fig. 3.4

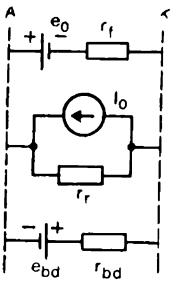


Fig. 3.5

not to differ markedly from the solution of the equation

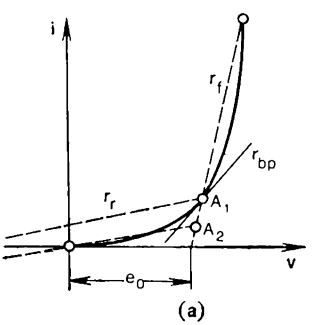
$$r_f/R_{eq} = R_{eq}/r_r$$

Therefore, if nothing is known about R_{eq}, the breakpoint must be located such that

$$r_{bp} \approx \sqrt{r_f r_r} \tag{3.2}$$

In the second case, the breakpoint is at position A_2 which is such that the reverse V/I characteristic passes through the origin, thereby rendering the diode model a passive one. The passive behaviour is essential in a number of applications for diode switches, notably when use is made of a capacitor charge. An applicable equivalent circuit appears in Fig. 3.6*b*. Because the forward resistance is small, the positions of points A_1 and A_2 on the voltage axis do not practically differ from the initial voltage e_0 which may be taken to give the cut-in threshold for the diode.

The breakpoint B is located (in building a model for the Zener diode) in a similar fashion.



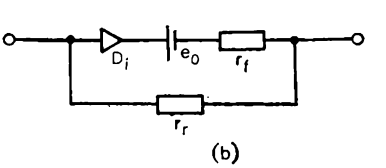


Fig. 3.6

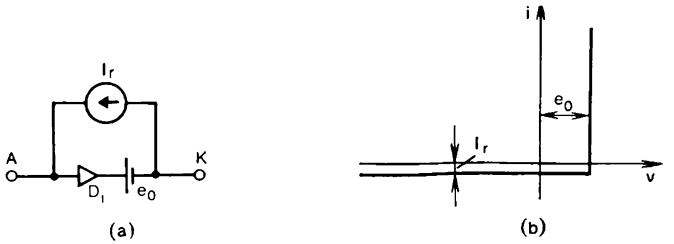


Fig. 3.7

In cases where the designer may neglect the dynamic resistance in the forward-bias region and the dynamic conductance in the reverse-bias region, a simplified piecewise-linear model may well serve its purpose (Fig. 3.7). In this model, the conducting diode is represented by a direct-voltage source e_0, and the nonconducting diode by a direct-current source I_r.

Nonlinear models. A distinction of such models is that the V/I characteristic of the diode is approximated by a function, $i = f(v)$, continuous over the entire range of operating voltages. Most frequently, recourse is had to an exponential model based on the V/I characteristic of an idealized p-n junction [34, 42, 43]

$$i_{p\text{-}n} = I_s \left[\exp\left(v_{p\text{-}n}/m\varphi_T\right) - 1\right] \tag{3.3}$$

where $i_{p\text{-}n}$ is the current across the junction, I_s is the reverse saturation current, φ_T is the thermal potential approximately equal to 26 mV at room temperature, m is a coefficient taking care of the discontinuities that exist in the transition region (for germanium, $m \approx 1$, and for silicon, $m \approx 1.2$ to 2), and $v_{p\text{-}n}$ is the potential difference between the p- and the n-region.

The characteristic of an idealized p-n junction (the dashed line in Fig. 3.2) runs precisely together with that of a real diode only within a limited portion near zero. The divergence of the forward portions of the characteristics is mainly due to the effect of the base bulk (or body) resistance r_b. If we allow for r_b, the forward V/I characteristic of a real diode can be deduced from Eq. (3.3) by substituting $v_{p\text{-}n} = v_f - ir_b$. I_s, which cannot be measured directly, can conveniently be expressed in terms of V_f measured at a given forward current I_f. Assuming that $I_f \gg I_s$, we obtain from Eq. (3.3)

$$I_s \approx I_f \exp\left[(I_f r_b - V_f)/m\varphi_T\right] \tag{3.4}$$

In the region of sufficiently high forward currents ($i_f \gg I_s$), the joint use of Eqs. (3.3) and (3.4) yields

$$v_f = V_f + m\varphi_T \ln\left(i_f/I_f\right) + (i_f - I_f)\, r_b \tag{3.5}$$

The dynamic forward resistance is given by

$$r_{f,d} = dv_f/di_f = r_b + m\varphi_T/i_f \tag{3.6}$$

Whether the first or the second term in Eq. (3.6) prevails will depend on the value of current.

The departure of the reverse characteristic of a real diode from the V/I curve of an idealized p-n junction is due to thermal (saturation) and leakage currents [14]. Their effect can be modelled by placing r'_r across the junction. In the circumstances, the reverse characteristic of the diode can be written

$$i = -I_s + v/r'_r + I_s \exp(v/m\varphi_T) \tag{3.7}$$

The last term is significant only at low values of v_r comparable with the thermal potential. The linear part of Eq. (3.7) is identical in form with the equation describing the reverse characteristic of the piecewise-linear model. For this reason, it is often arranged so that $r_r = r'_r$ and $I_0 = I_s$.

Applicability of static models. Ideal-diode models are usually invoked to represent the performance of a pulse switch at a design stage where accuracy is of minor importance, and also when the transient response(s) of a pulse circuit are examined only approximately. The piecewise-linear model is resorted to when the dynamic resistance of the switching element defines some functional parameter of the circuit (such as gain, transfer function, noise amplitude, etc.). If we are interested in an approximate analysis of operating conditions (with a margin of safety) and performance of a switch, recourse is had to the simplified model. The exponential model can be used to advantage in computer-aided circuit design and also in cases where the functional parameter of the switch depends on the rate of change in the dynamic resistance of the diode upon switching, for example, in evaluating the minimum amplitude of the switching signal.

Identification of static models. The parameters of the static diode models examined above are found from the values given in data sheets or, if necessary, by experiment.

For most applications, the forward V-I characteristic is given by specifying the maximum forward voltage $V_{f,\max}$ at a specified forward current I_f (Table 3.1). Data sheets also define the area encompassing 95% spread in the forward V-I characteristic (Fig. 3.8).

By substituting for I_f, V_f, i_f, and v_f in Eq. (3.5) the coordinates of two arbitrary points on the boundary curves, we can readily find the bulk base resistance r_b

$$r_b = \frac{V_{f1} - V_{f2} - m\varphi_T \ln(I_{f1}/I_{f2})}{I_{f1} - I_{f2}}$$

Points *1* and *2* in Fig. 3.8 represent $r_{b,\max}$, and points *1'* and *2'*, $r_{b,\min}$. In most cases, the designer is interested in the maximum

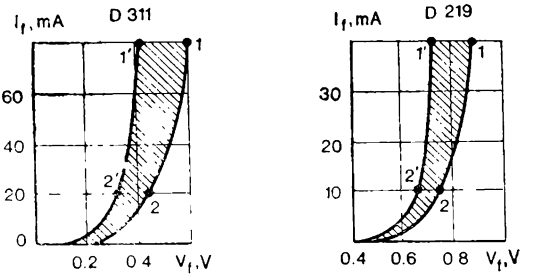


Fig. 3.8

permissible value of V_f. It can be found by substituting $V_{f,\,max}$ for V_f and setting r_b equal to $r_{b,\,max}$ in Eq. (3.5).

The parameters of piecewise-linear models are evaluated on the basis of the maximum forward current $I_{f,\,max}$ at a specified temperature. The corresponding value of voltage, $V_{f,\,max}$, is found from Eq. (3.5). The forward resistance r_f and the initial voltage e_0 are found by solving jointly Eqs. (3.2), (3.5) and (3.6), complemented with the equality $r_f I_{f,\,max} = V_{f,\,max} - e_0$.

The model parameters of typical pulse diodes for two values of $I_{f,\,max}$ are given in Table 3.1.

Table 3.1. Parameters of the Forward V/I Characteristic

Diode type	Data sheet values		$I_{f,max}$, mA	Model parameters			
	I_f, mA	V_f, V		r_b, Ω	$V_{f,max}$, V	r_f, Ω	e_0, V
Germanium, Д311	10	0.4	1	2.3	0.23	52	0.22
			10		0.4	7.3	0.24
Silicon, Д219	50	1	1	2.9	0.75	63	0.57
			10		0.84	8.9	0.73

The reverse V/I characteristic of a diode is defined by specifying the maximum reverse current, $I_{r,\,max}$, at a specified V_r, usually equal to the maximum value (Table 3.2). Then, $r_r \approx V_r/I_{r,\,max}$. The reverse saturation current I_s is found from Eq. (3.4) where I_f and V_f are replaced by the coordinates of an arbitrary point on the minimum boundary curve (say, point *2* in Fig. 3.8). Then the value of I_s thus found will be a maximum one.

The forward voltage across the diode has a negative temperature coefficient (2 to 4 mV/°C). The reverse saturation current rises ex-

Table 3.2. Parameters of the Reverse V/I Characteristic

Diode type	Data sheet values		Model parameters	
	V_r, V	$I_{r,\,max}$, μA	I_s, μA	r_r, MΩ
Д311	30	100	0.44	0.3
Д219	70	1	5×10^{-7}	70

onentially with increasing temperature. Approximately, it may be assumed that it almost doubles in magnitude for every ten degrees of rise in temperature. More accurately, the temperature dependence can be deduced from the values given in the data sheet.

Dynamic Models of Crystal Diodes

A diode used as a switch shows a time lag in its response to fast-changing input signals because of two factors, namely: the re-charging of the barrier capacitance C_b of the p-n junction, and the change in the charge Q associated with the carriers injected into the base region.

The time it takes a diode switch to go to the ON state is the sum of two terms. The first is called the *turn-on delay time*; the p-n junction is cut off so that $Q = 0$, and the time lag in the response is solely due to C_b. The effect of this term is illustrated by the equivalent circuit in Fig. 3.9*a* where D is a lag-free diode which can be characterized by the static equivalent circuit fitting the reverse V-I characteristic. The turn-on delay time is complete when the capacitor voltage becomes equal to the cut-in voltage e_0 of the diode. The second term is the *transition time*; during this interval the base accumulates a charge Q so that the base resistance decreases towards a steady-state value. The sum of the two intervals is known as the *forward recovery time*. Immediately after a turn-on, the voltage across the diode (the forward pulse voltage drop) exceeds the steady-state value. If, however, the forward current is not over 10 mA, this excess will be negligible.

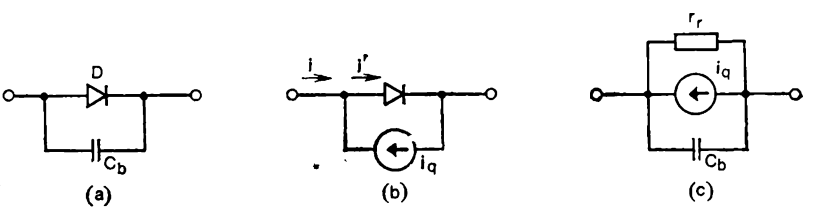


Fig. 3.9

The time it takes a diode to turn off is likewise the sum of two terms, the storage time and the transition time, both giving that is termed as the *reverse recovery time*, τ_{rr}. The storage time refers to the interval required for the stored minority charge Q to become zero, and the transition time refers to the interval which is required for the charge density q to re-distribute in the base region. The distribution of q in the x-direction orthogonal to the p-n junction plane for the various intervals of the turn-off time is illustrated in Fig. 3.10. Curve *1* applies to the storage interval. The area bounded by curve *1* and the coordinate axes is numerically equal to the stored charge. The density q_0 is a maximum in the region adjacent to the emitter ($x = 0$). As always occurs in diffusion, the flow of injected carriers determining the current through the junction is directed from the region of higher concentration towards the region of lower concentration, that is, in the positive x-direction. The value of q_0 is uniquely connected to the voltage across the junction by a relation of the form

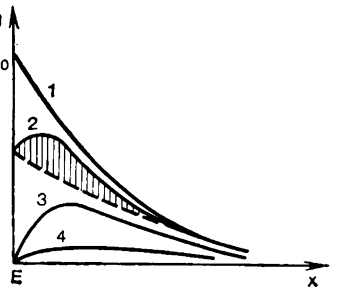


Fig. 3.10

$$q_0 = ai'(v) \tag{3.8}$$

where $i'(v)$ is the V/I characteristic of the junction and a is the scale factor.

When a diode turns off, the flow of current across the junction is reversed. Now the charge is swept from base to emitter instead of from emitter to base. The distribution $q(x)$ takes the form represented by curve *2* in Fig. 3.10. The storage interval lasts so long as $q_0 > 0$, and the diode remains conducting because the voltage v lies within a high-slope portion of the V/I curve in accordance with Eq. (3.8).

The instant of cutoff at which q_0 vanishes is represented by curve *3*. This instant marks the beginning of the transition interval accompanied by a simultaneous change in the voltage across and the current through the diode.

During the storage interval, a switching diode can be represented by the equivalent circuit of Fig. 3.9*b* where D is a lag-free diode answering the static equivalent circuit for the low-resistance portion of the V/I characteristic. The current i' is equal to the steady-state forward current through the diode at a given q_0. This current corresponds to the fraction of Q lying under the dashed exponential curve in Fig. 3.10. The current source i_q simulates the current component due to the shaded area under curve *2*. The value of i during turn-off

can be written as a difference

$$i = i(0) - i_{\sim}(t)$$

where $i(0)$ is the initial steady-state value which is the same as $i'(0)$, and $i_{\sim}(t)$ is the change caused by the turn-off signal applied to the diode.

Under the Laplace transformation, i_q and $i_{\sim}$ are connected by a relation of the form

$$i_q(s) = i_{\sim}(s) K(s)$$

where $K(s)$ is the transfer function specified by the transcendental function of the complex-frequency variable s. This function can be evaluated by solving the equations describing the diffusion and drift of injected carriers in the base region. Its form is decided by the geometry and electrophysical properties of the base region. Since in the steady state, $i_q = 0$, so the transcendental function $K(s)$ must satisfy the condition $K(0) = 0$.

For rough calculations, it will suffice to approximate $K(s)$ by a bilinear function. Then,

$$i_q(s) \approx i_{\sim}(s)\, s\tau_1/(1 + s\tau_2) \tag{3.9}$$

The time constants τ_1 and τ_2 are decided by the mean life times of the injected carriers, τ. For alloyed diodes, $\tau_1 \approx 0.5\tau$ and $\tau_2 \approx 0.75\tau$. For retarding-field (diffused-base) diodes, $\tau_1 = \tau/(1 + \gamma_E^2)$ and $\tau_2 = \tau$, where $\gamma_E = aL/N$ is the field factor, a is the concentration (number density) gradient, N is the impurity concentration, and L is the diffusion length of the injected carriers.

In some cases (such as when the calculations are done on a computer), it will be more convenient to write Eq. (3.9) as a differential equation, so that one can go over from $i_{\sim}$ to the true value of i,

$$\tau_2 (di_q/dt) + i_q = \tau_1 (di_{\sim}/dt) = -\tau_1 (di/dt) \tag{3.10}$$

The condition for a diode to be cut off, $i_q = i_{q,\,\mathrm{bp}} = -i$, defines the instant at which the storage interval ceases and the transition interval commences.

Consider the case where a diode is turned off by a direct current, that is, $i = -I_1 =$ constant. Denoting the initial forward current as I_f, we obtain

$$i_{\sim}(s) = (I_f + I_1)/s$$

whence

$$i_q(s) = \frac{(I_f + I_1)\,\tau_1}{1 + s\tau_2}$$

$$i_q(t) = \frac{(I_f + I_1)\,\tau_1}{\tau_2} \exp(-t/\tau_2)$$

Using the cutoff condition, $i_q(\tau_s) = i_{q,\,bp} = I_1$, we find that

$$\tau_s = \tau_2 \ln \frac{(1 + I_f/I_1)\tau_1}{\tau_2} \tag{3.11}$$

With alloyed diodes for which $\tau_1/\tau_2 \approx 2/3$, Eq. (3.11) yields sufficiently accurate results for $I_1 < 2I_f$. At $I_1 = 2I_f$, we obtain from Eq. (3.11) that $\tau_s = 0$, which contradicts experiment. Relationships that will hold at $I_1 > 2I_f$ require a more rigorous approximation of $K(s)$. Other methods for estimating τ_s for $I_1 =$ constant can be found in [14]. With diffused-base diodes Eq. (3.11) yields sufficiently accurate results over the entire range of permissible turn-off currents.

To build a diode model for the transition interval, we shall first consider the case where the voltage across the diode remains constant after cutoff. In the circumstances, the current through the diode is solely due to the stored minority carrier charge which decays along curve *3* (or *4*) in Fig. 3.10, that is, $i = -i_q$. Substituting this equality in Eq. (3.10) gives

$$(\tau_2 - \tau_1)(di_q/dt) + i_q = 0$$

Solving the above equation subject to the initial conditions gives

$$\begin{aligned} i_q(t) &= i_{q,\,bp} \exp\left(-\frac{t}{\tau_2 - \tau_1}\right) \\ i_q(s) &= i_{q,\,bp} \frac{\tau_2 - \tau_1}{1 + s(\tau_2 - \tau_1)} \end{aligned} \tag{3.12}$$

In the general case, when the cutoff voltage is reversed, the total current through the diode contains, in addition to i_q, also the current required to re-charge C_b and the reverse current due to r_r. Because these currents are not related to the flow of injected carriers in the base, the validity of Eq. (3.12) is not violated. An equivalent circuit for a switching diode during the transition interval is shown in Fig. 3.9*c*.

For diffused-base diodes, the exponential time constant in Eq. (3.12) is $\tau_2 - \tau_1 = \tau/\gamma_E^2$. For charge-storage diodes, special measures are taken to enhance the retarding field ($\gamma_E \geqslant 5$ to 7), so that the fall time for the reverse current is a small fraction of the storage time. This property of charge-storage diodes is utilized in diode pulse amplifiers and shapers [56].

Identification of parameters for a dynamic model. The parameters of a dynamic model can be identified from data sheets which usually state the maximum reverse recovery time t_{rr} at specified values of $I_1 \approx I_f$. The instant marking the end of the reverse recovery interval depends on the specified value of reverse current $i_q = i_{q,\,max}$. Then, in accordance with Eqs. (3.11) and (3.12), for an alloyed diode

$$\tau = 4t_{rr}/[1 + \ln(I_f/i_{q,\,max})]$$

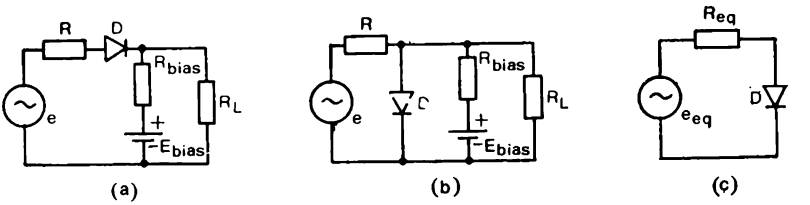


Fig. 3.11

For diffused-base diode, the fall time of the reverse current i_q may be neglected; then, $\tau \approx 1.6t_{rr}$.

The value of barrier capacitance is a function of voltage, so data sheets usually state it for some specified voltage, V_{nom}. For an arbitrary value of V_1, the data-sheet (nominal) barrier capacitance is

$$C_{b,1} = C_{nom} \sqrt[n]{(V_{nom} + \varphi_0)/(V_1 + \varphi_0)}$$

where φ_0 is the contact potential difference (for germanium diodes it is about 0.4 V, and for silicon diodes, 0.6 to 0.8 V), and n is a factor (equal to 2 for abrupt junctions, and to 3 for graded junctions).

Piecewise-linear models ordinarily use the value of barrier capacitance averaged over the interval from V_1 to V_2:

$$\bar{C}_b = C_{b,1} \frac{1 - [(V_2 + \varphi_0)/(V_1 + \varphi_0)]^{(n-1)/n}}{[(n-1)/n][1 - (V_2 + \varphi_0)/(V_1 + \varphi_0)]}$$

Practical Diode Switches

Let us examine two typical diode switch configurations differing in how the diode is connected in the circuit. In one configuration, it is series-connected (Fig. 3.11*a*), and in the other, it is parallel-connected (Fig. 3.11*b*); In either case, the control circuit contains a source of command signal e and a resistance R which includes the source resistance. The source of bias voltage E_{bias} = constant, via a bias resistor R_{bias}, serves to set the desired threshold voltage for the diode switch. The load resistance is R_L. As an alternative, the diode D can be connected in the reverse direction. By applying Thevenin's theorem to the linear portion of either circuit, it can be transformed to the circuit of Fig. 3.11*c* where:

—for the series configuration

$$\begin{aligned} R_{eq} &= R + R_{bias}R_L/(R_{bias} + R_L) \\ e_{eq} &= e - E_{bias}R_L/(R_L + R_{bias}) \end{aligned} \tag{3.13}$$

—for the parallel configuration

$$\begin{aligned} R_{eq} &= (1/R + 1/R_{bias} + 1/R_L)^{-1} \\ e_{eq} &= R_{eq}(e/R + E_{bias}/R_{bias}) \end{aligned} \tag{3.14}$$

The threshold value of control voltage, e_{th}, depends on when the diode is turned on. The equivalent circuit of Fig. 3.6*b* enables us to find the voltage across the non-conducting diode:

$$v = e_{eq} r_r / (R_{eq} + r_r)$$

Hence, the equivalent threshold voltage is

$$e_{eq,\,th} = e_0 (1 + R_{eq}/r_r) \tag{3.15}$$

Substituting Eq. (3.15) into Eqs. (3.13) and (3.14) and solving them for e yields:

—for the series configuration

$$e_{th} = e_0 (1 + R_{eq}/r_r) + E_{bias} R_L/(R_L + R_{bias})$$

—for the parallel configuration

$$e_{th} = (R/R_{eq})\, e_0 (1 + R_{eq}/r_r) - E_{bias} R/R_{bias}$$

The above equations allow for the effect of the diode parameters on the operating threshold and evaluate its stability.

Transient response of a diode switch. Suppose that the diode in the circuit of Fig. 3.12*a* is switched by an ideal rectangular pulse and that the time lag is decided by the diode and the shunt stray

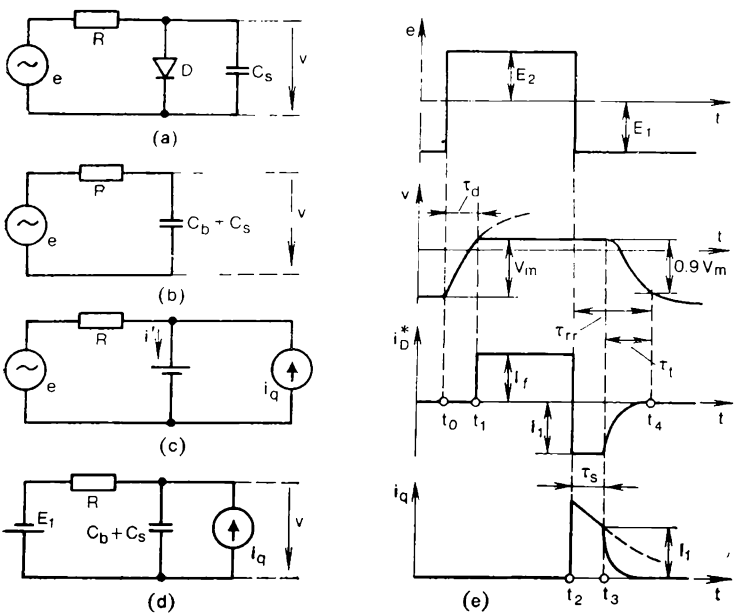


Fig. 3.12

capacitance C_s. The time records of circuit operation are shown in Fig. 3.12*e*. The diode current constituted by the flow of minority carriers across the junction is denoted by i_D^*. A complete switching cycle consists of a turn-on delay interval or time ($t_0 < t < t_1$), a steady ON state ($t_1 < t < t_2$), a turn-off delay interval or storage time ($t_2 < t < t_3$), and a transition interval or time ($t_3 < t < t_4$). The equivalent circuits for the various intervals appear in Fig. 3.12*b* through *d* and correspond to the dynamic equivalent circuits for a diode in Fig. 3.9. The static model to use will be a simplified piecewise-linear model for $I_r \approx 0$, in which i_D^* is shown to change abruptly at turn-on, t_1.

The turn-on delay is associated with the time required to re-charge $C_{eq} = C_s + \bar{C}_b$ (see Fig. 3.12*b*). The voltage on the diode varies exponentially with a time constant $\tau_{eq} = RC_{eq}$:

$$v = E_1 - (E_1 + E_2) \exp(-t/\tau_{eq})$$

The turn-on delay interval is completed at time t_2 when $v = e_0$, so

$$\tau_d = \tau_{eq} \ln [1 + (E_2 + e_0)/(E_1 - e_0)]$$

According to the equivalent circuit of Fig. 3.12*c* which depicts the diode switch in the steady ON state for $i_q = 0$, we may write

$$i' = I_f = (E_1 - e_0)/R$$

During the storage interval, the turn-off current remains constant, $I_1 = (E_2 + e_0)/R$, and τ_s is given by Eq. (3.11).

The transition time τ_t for a diode switch can be found using the equivalent circuit of Fig. 3.12*d* and Eq. (3.12). Thus, the diode voltage is a sum of two exponential terms:

$$v = \frac{(E_2 + e_0)\{\tau_{eq} \exp(-t/\tau_{eq}) - (\tau_2 - \tau_1) \exp[-t/(\tau_2 - \tau_1)]\}}{\tau_{eq} - (\tau_2 - \tau_1)} - E_2$$

where $\tau_{eq} = R(C_s + \bar{C}_b)$.

Assuming that at time t_4 the output signal is 10% of its steady-state amplitude and that $(\tau_2 - \tau_1) < \tau_{eq}$, we obtain

$$\tau_t \approx 2.3\tau_{eq} + \frac{\tau_2 - \tau_1}{1 - (\tau_2 - \tau_1)/\tau_{eq}} \tag{3.16}$$

In most cases, especially with diffused-base diodes, the second term in Eq. (3.16) is negligibly small. On the whole, it is to be noted that owing to the low value of barrier capacitance ($C_b \approx 1$ to 3 pF) and the short life time of injected carriers, present-day diodes may be treated as lag-free elements and represented by a static model.*

* Here and elsewhere, the term static refers to both d.c. (static) voltages and currents, and to quasi-static ones varying at a rate so small that the time lag (inertia) of the switching element may be ignored.

3.3. TRANSISTOR SWITCHES

Transistor switches are key elements in the more sophisticated pulse circuits, such as logic gates, flip-flops, etc. We shall consider bipolar-transistor models used in the analysis of switches, and also the working principles and responses of the common-emitter (CE) transistor switch as it is basic to most pulse circuits.

Static Models of the Bipolar Transistor

A transistor (Fig. 3.13*a*) is a system of two *p-n* junctions and a common base region. In contrast to an isolated diode, the reverse current across each junction may have an additional component, i_q, due to the minority carriers injected into the base through the other junction. Therefore, in an equivalent circuit diagram, a bipolar transistor can be depicted as a connection of two diodes, D_E and D_C, simulating the V/I responses of the emitter and collector junctions of the transistor, and sources of currents i_{qE} and i_{qC}, simulating the interaction between the junctions. Figure 3.13*b* shows an equivalent circuit for an *n-p-n* transistor; in the case of a *p-n-p* transistor, the

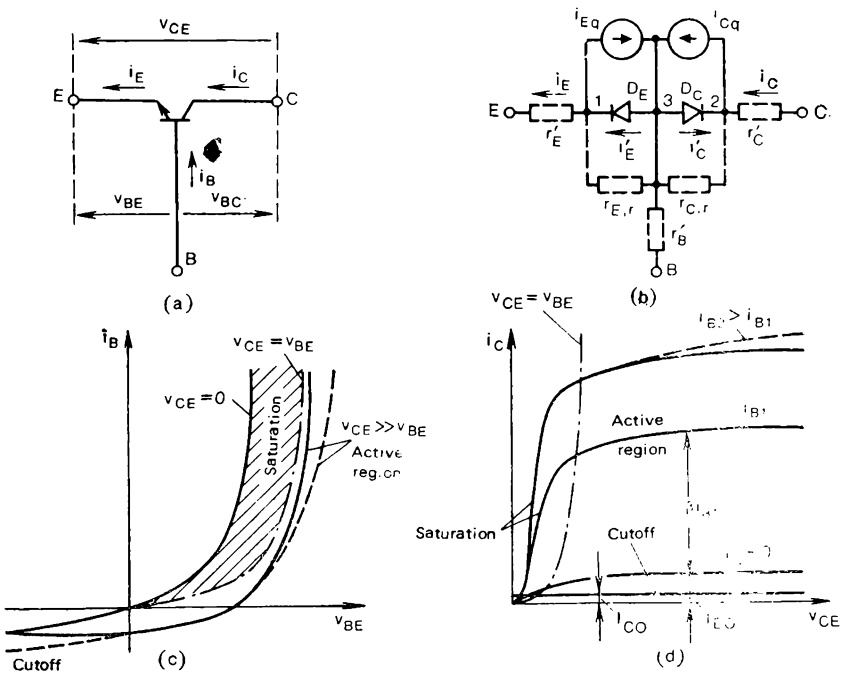


Fig. 3.13

diodes must be turned over.

In the static regime, i_{qE} and i_{qC} are defined as

$$\begin{aligned} i_{qC} &= \alpha i'_E \\ i_{qE} &= \alpha_I i'_C \end{aligned} \tag{3.17}$$

where i'_E and i'_C are the injection currents in the emitter and collector junctions, α and α_I are the current gain from emitter to collector and from collector to emitter, respectively. Transistors are built so that α is unity very closely (as a rule, $1 - \alpha < 0.05$ to 0.02), and α_I has a small value ($1 - \alpha_I$ is less than 0.1 to 0.2). The value of α_I is especially small (of the order of 0.01 to 0.15) for drift transistors where the flow of carriers from collector to emitter opposes the internal field.

Transistor models can be classed according to how they approximate the *V-I* characteristics of the emitter and collector junctions.

Exponential (Ebers-Moll) model of the transistor. This model is based on approximating the *V-I* characteristics of the junctions by the relation defined in Eq. (3.3). According to Eq. (3.3) and Kirchhoff's current law for nodes *1* through *3* of the circuit in Fig. 3.13*b*, the emitter, collector and base currents are defined as

$$\begin{aligned} i_E &= I_{sE}\left(\exp\frac{v_{BE}}{m\varphi_T} - 1\right) - \alpha_I I_{sC}\left(\exp\frac{v_{BC}}{m\varphi_T} - 1\right) \\ i_C &= \alpha I_{sE}\left(\exp\frac{v_{BE}}{m\varphi_T} - 1\right) - I_{sC}\left(\exp\frac{v_{BC}}{m\varphi_T} - 1\right) \\ i_B &= (1-\alpha) I_{sE}\left(\exp\frac{v_{BE}}{m\varphi_T} - 1\right) \\ &\quad + (1-\alpha_I) I_{sC}\left(\exp\frac{v_{BC}}{m\varphi_T} - 1\right) \end{aligned} \tag{3.18}$$

where I_{sE} and I_{sC} are the reverse saturation currents of the emitter and collector junctions, v_{BE} and v_{BC} are the base-to-emitter and base-to-collector voltage, respectively (see Fig. 3.13*a*).

Equations (3.18) describe the behaviour of a transistor in any region: active, saturation and cutoff (except breakdown and punch-through). It appears useful to supplement them by a relation known from transistor theory:

$$\alpha I_{sE} \approx \alpha_I I_{sC} \tag{3.19}$$

The characteristics of real and ideal transistors differ mainly because the models ignore variations in α and α_I with the region in which a given transistor is operating. To a certain extent, they are also affected by the reverse resistances r_{rC} and r_{rE}, and the bulk resistances of the emitter, collector and base regions, r'_E, r'_C and r'_B, shown in the plot of Fig. 3.13*b* by the dashed line.

The general equations (3.18) and (3.19) can be re-arranged to derive partial equations describing the behaviour of a transistor in the various regions.

The reverse collector (emitter) current is the current through the collector (emitter) junction at a specified inverse voltage V_{CB} (V_{EB}), with the emitter (collector) circuit open. Let the respective reverse currents of an idealized transistor be denoted as I_{CO} and I_{EO}. With a sufficiently high cutoff voltage (V_{CB}, $V_{EB} \gg m\varphi_T$), we may neglect the exponential terms in Eq. (3.18). Then,

$$I_{CO} = (1 - \alpha\alpha_I)\, I_{sC}$$
$$I_{EO} = (1 - \alpha\alpha_I)\, I_{sE}$$

The reverse currents of a real transistor, with allowance for r_{rE} and r_{rC}, are given by

$$I_{CBO} = I_{CO} + |\, V_{CB}\, |/r_{rC}$$
$$I_{EBO} = -(I_{EO} + |\, V_{EB}\, |/r_{rE})$$

The above relations can be used to identify r_{rC} and r_{rE} from the data-sheet values of I_{CBO}, V_{CB}, I_{EBO}, and V_{EB}.

A transistor is driven far beyond cutoff when it is reverse-biased at both junctions. Then, using Eqs. (3.18) and (3.19), we obtain

$$i_E = -I_{sE} + \alpha_I I_{sC} \approx -I_{sE}\,(1 - \alpha)$$
$$i_C = -I_{sE}\,\alpha + I_{sC} \approx I_{sC}\,(1 - \alpha_I)$$

Since α tends to unity, we may set $i_E \approx 0$, $i_C \approx I_{CO}$, and $i_B = i_E - i_C = -I_{CO}$.

If we include r_{rE} and r_{rC}, the reverse currents far beyond cutoff can be written

$$\begin{aligned} i_c &= I_{CBO} \\ i_e &= -V_{CE}/r_{rE} \\ i_b &= -(I_{CBO} + V_{EB}/r_{rE}) \end{aligned} \qquad (3.20)$$

Equations (3.20) evaluate the parameters of a switching transistor in the OFF state.

When an *n-p-n* transistor is forward-biased, then $V_{CE} > 0$. That is, in a conducting transistor, the emitter junction is always forward-biased. On setting $i_E \gg I_{EO}$, the equations describing the V/I characteristic of a conducting transistor can be written as

$$v_{BE} = m\varphi_T \ln (i_C/I_{sE}) \left[\frac{1}{\alpha} + \frac{K_{sat} - 1}{\beta\,(1 - \alpha\alpha_I)}\right] \qquad (3.21)$$

$$i_C = \beta i_B + I_{CEO}\left(1 - \exp\frac{v_{BE} - v_{CE}}{m\varphi_T}\right) \qquad (3.22)$$

where $\beta = \alpha/(1 - \alpha)$ is the current transfer ratio (gain) of the base, $K_{sat} = \beta i_B/i_C$ is the saturation factor, and $I_{CEO} = (1 + \beta) I_{CBO}$ is the collector reverse current with the base open.

The transistor is operating in the active region when $v_{CE} > v_{BE}$ (see Fig. 3.13*c* and *d*). In this region, the exponential term in Eq. (3.22) is small, and the transistor has a high output impedance on the collector side. In the saturation region, when $v_{CE} < v_{BE}$, the exponential term increases in absolute value, and the output impedance is abruptly reduced. It is customary to consider a boundary region, with $v_{CE} = v_{BE}$, where, in accordance with Eq. (3.22), $i_C = \beta i_B$ and, as a consequence, $K_{sat} = 1$. As the transistor moves into the saturation region, i_C decreases in comparison with βi_B, so that the condition for saturation can be written as

$$K_{sat} > 1 \text{ or } \beta i_B > i_C \tag{3.23}$$

In the active region, $K_{sat} = 1 - I_{CEO}/i_C$ approaches unity with increasing collector current. If we write I_{sE} in terms of the nominal base-to-emitter voltage $V_{BE,\,nom}$, as measured at a specified value of $I_{E,\,nom}$ in the active region, we may re-write Eq. (3.21) as

$$v_{BE} = V_{BE,\,nom} + m\varphi_T \ln \left\{ (i_C/I_{E,\,nom}) \left[\frac{1}{\alpha} + \frac{K_{sat} - 1}{\beta (1 - \alpha\alpha_I)} \right] \right\} \tag{3.24}$$

If the data sheet specifies tha base-to-emitter voltage in the saturation region, then

$$V_{BE,\,nom} = V_{BE(sat)} + m\varphi_T \ln \left[1 + \frac{K_{sat,\,nom} - 1}{(1 + \beta)(1 - \alpha\alpha_I)} \right]$$

where $K_{sat,\,nom}$ is the saturation factor at which $V_{BE\,(sat)}$ is measured. To adjust for the effect of the base bulk resistance, Eq. (3.24) can be extended to include a further term, $[i_B - I_{E,\,nom}/(1 + \beta)]\, r'_B$.

The collector-to-emitter voltage in the saturated region can be found by simultaneously solving Eqs. (3.21) and (3.22) for v_{CE}. On setting $i_C = I_{C(sat)}$, we have

$$V_{CE(sat)} = m\varphi_T \ln \frac{1}{\alpha_I} \frac{K_{sat} + \beta (1 - \alpha_I)}{K_{sat} - 1 + I_{CEO}/I_{C(sat)}} \tag{3.25}$$

If $K_{sat} \ll \beta (1 - \alpha_I)$ and $I_{CEO}/I_{C(sat)} \ll K_{sat} - 1$, then Eq. (3.25) can be simplified as follows:

$$V_{CE(sat)} \approx m\varphi_T \left[\ln \frac{\beta}{\beta_I} - \ln (K_{sat} - 1) \right] \tag{3.26}$$

where $\beta_I = \alpha_I/(1 - \alpha_I)$ is the reverse-bias (inverse) current transfer ratio (gain) of the base. Equations (3.25) and (3.26) can be extended to include a further term, $I_{C(sat)}\, r'_C$, to allow for the voltage drop across the bulk resistance of the collector. The effect of r'_C is appreciable in planar (IC) transistors with a lightly doped collector region.

Equations (3.24) through (3.26) make it possible to evaluate the residual voltages between the terminals of a switching transistor in the ON state.

Owing to the low value of $V_{CE(sat)}$, it appears advantageous to use the collector and emitter terminals as the output port of a switching transistor. With such an arrangement, the switching performance is markedly better than it is with an isolated *p-n* junction.

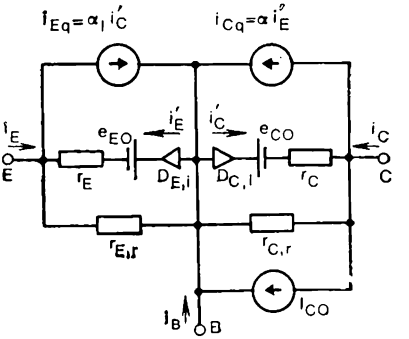


Fig. 3.14

A switching transistor can also be reverse-biased, in which case $v_{CE} < 0$, and the emitter and collector junctions interchange functions. Appropriate design relations can be derived from Eqs. (3.21) through (3.26), on interchanging the forward and reverse parameters. As often as not, a switching transistor can alternate between forward- and reverse-biased operation depending on which signal is applied to the input.

Piecewise-linear models. These models are based on a piecewise-linear approximation of the *V/I* curves of the emitter *p-n* junction, D_E, and the collector *p-n* junction, D_C, in the equivalent circuit of Fig. 3.13*b*. If the reverse *V/I* characteristics of D_E and D_C are such that $i_{E,r} = I_{sE}$ and $i_{C,r} = I_{sC}$, and the junctions are open, the transistor will operate in a regime fully identical with the deep cutoff region of the exponential model described by Eq. (3.20). About the same result can be obtained by using the equivalent circuit of Fig. 3.14 in which, with the junctions open, $i_{E,r} = i_{C,r} = 0$, and the initial reverse currents are simulated by a current generator, I_{CO}.

Like the exponential (Ebers-Moll) model, the piecewise-linear model illustrated by the equivalent circuit of Fig. 3.14 can be used to represent the transistor in any one of the operating regions.

A simplified piecewise-linear model can be derived from the equivalent circuit of Fig. 3.14 on setting $r_E = r_C = 0$, $r_{r,E} = r_{r,C} = \infty$, and $I_{CO} = I_{CBO}$, where I_{CBO} is the reverse collector current at a specified (usually, maximum) collector-to-base voltage, with the emitter open-circuited. With this model, the transistor junctions in the ON state are simulated by a voltage generator, and in the OFF state by a current generator (Fig. 3.15*a*). Its use is justified if the resistance between the transistor terminals is a small fraction of the equivalent resistance of the external circuit when the junctions are in the ON (conducting) state, and many times its value when the junctions are in the OFF (nonconducting) state. The input and output

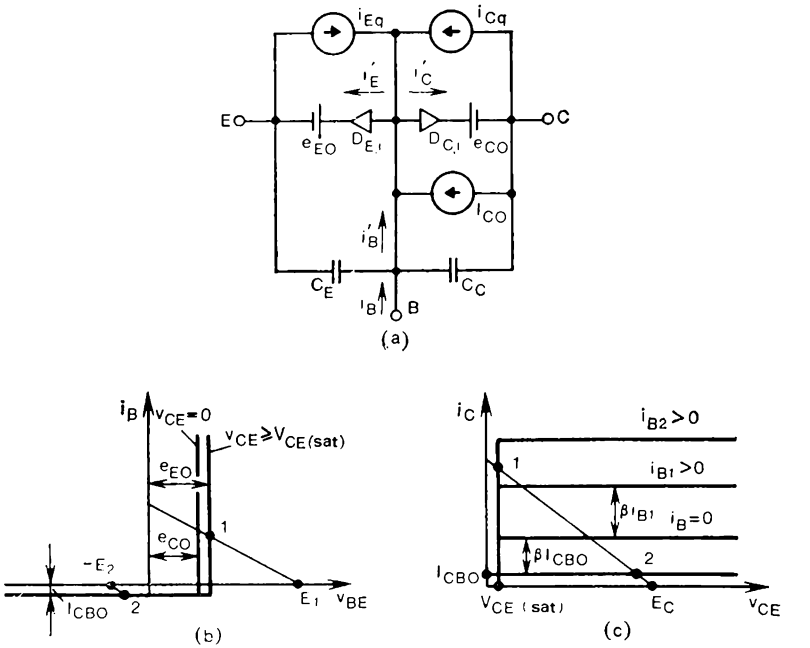


Fig. 3.15

characteristics of the simplified piecewise-linear model for a forward-biased transistor are shown in Fig. 3.15*b* and *c*.

The simplest model of a switching transistor is an ideal transistor in which the junctions are simulated by ideal diodes. Its characteristics are the same as those of the piecewise-linear model for $I_{CBO} = 0$ and $e_{EO} = e_{CO} = 0$.

Identification of parameters for static models. The values of α and β are related to the data-sheet value of the small-signal forward current transfer ratio h_{fe} which is defined as the ratio of a change in output current to the change in input current that causes it, as measured for a specified average collector current with the output short-circuited for a.c.

In most cases, data sheets specify the base current as input, and the collector current as output. In the circumstances, it may be taken, at least approximately, that $\beta = h_{fe}$ and $\alpha = h_{fe}/(1 + h_{fe})$. Sometimes, data sheets specify the static forward current transfer ratio for the common-emitter configuration, h_{FE}, defined as the ratio of the direct collector current to the direct base current, given V_{CE} and I_E. This ratio is close to β only at small values of I_{CBO} in comparison with I_E and in cases where h_{FE} only slightly depends

on the operating conditions. For approximate calculations, however, it will be safe to take $\beta \approx h_{FE}$.

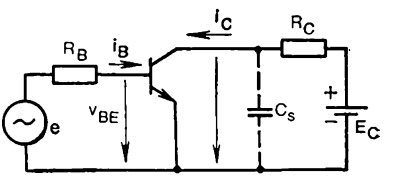


Fig. 3.16

The usual procedure for determining α_I is to measure $V_{CE(sat)}$ with the collector open-circuited ($I_{C(sat)} = 0$) and to use Eq. (3.25) in which K_{sat} is set equal to infinity. The parameters of the V/I characteristics for the p-n junctions are identified, using techniques similar to those employed for the static models of diodes. The typical values of transistor model parameters are $\alpha = 0.95$ to 0.99, $\alpha_I = 0.1$ to 0.5, and $e_{EO} = e_{CO} = 0.6$ to 0.8 V for silicon, and $e_{EO} = e_{CO} = 0.25$ to 0.3 V for germanium.

Temperature stability. An important factor for transistors, apart from the parameter stability of the p-n junctions, is the temperature drift in β and β_I, stated in terms of the positive temperature coefficient whose value ranges from 0.005 to 0.01 β_n/°C, where β_n is the transfer ratio (gain) at normal temperature.

Common-emitter transistor switch. The transistor switch shown in Fig. 3.16 has a switched and a switching (control) circuit. The switched circuit contains a resistor R_C and a collector power supply E_C, both placed between collector and emitter. Under steady-state conditions, v_{CE} is connected to i_C by Kirchhoff's voltage (second) law:

$$v_{CE} = E_C - i_C R_C \tag{3.27}$$

On the output characteristics of the simplified transistor model (see Fig. 3.15*c*), this equation is represented by a load line (*1*, *2*).

The switched circuit is completed when the transistor is at saturation (point *1*). Then, according to Eq. (3.27),

$$i_C = I_{C(sat)} = (E_C - V_{CE(sat)})/R_C$$

For the transistor to operate in the switching mode, it is presumed that $V_{CE(sat)} \ll E_C$, so that we may set

$$I_{C(sat)} \approx I_{sc} = E_C/R_C$$

The switched circuit is open when the transistor is at cutoff (point *2*). Then, according to Eq. (3.27),

$$v_{CE} = E_C - I_{CBO} R_C$$

For the transistor to operate in the switching mode, it is further presumed that $I_{CBO} R_C \ll E_C$, so that we may state that $v_{CE} \approx E_C$.

The control circuit of a transistor switch is formed by a resistor R_B and a supply of control voltage e, both connected across the emitter

junction of the transistor. With this arrangement, the emitter lead is common to both the switched and the control circuits. Under steady-state conditions, v_{BE} and i_B satisfy Kirchhoff's voltage law:

$$v_{BE} = e - i_B R_B \tag{3.28}$$

Consider the operation of the transistor switch for two values of control voltage, E_1 and E_2, assuming that the former drives the transistor into conduction and the latter, into cutoff. With $e = E_1$, Eq. (3.28) yields on the input characteristic of the simplified transistor model (see Fig. 3.15*b*) a straight line passing through point *1*. In the circumstances,

$$v_{BE} = e_{EO}$$

and, according to Eq. (3.28),

$$i_B = (E_1 - e_{EO})/R_B$$

To enhance the reliability of a transistor switch, it is customary to choose the minimal saturation factor (of the order of 1.5 to 2) that must be maintained for the worst combination of transistor parameters. On setting $i_B = K_{sat,min} I_{C(sat)}/\beta$ (according to the definition of the saturation factor), we can find the minimum value of E_1 required for the switch to operate (close) reliably

$$E_{1,\,min} = e_{EO,\,max} + R_B K_{sat,\,min} I_{C(sat)}/\beta_{min}$$

With $e = E_2$, Eq. (3.28) yields the input characteristic of Fig. 3.15*b* a straight line passing through point *2*. In this case, it follows from Eq. (3.28) that

$$v_{BE} = E_2 + I_{CBO} R_B$$

Within the framework of the simplified model in question, the cut-in (enable) threshold for the transistor is e_{EO}. In practice, as a way of enhancing the performance of the transistor switch, the maximum voltage on the base of the OFF transistor must lie below the threshold value by an amount ΔV_N to provide a margin against an enabling noise. The choice of ΔV_N varies with the function and operating conditions of a switch and usually ranges from 0.3 to 1.5 V. (An excessive increase in ΔV_N would reduce the speed, or frequency response, of the switch). It may be added that the noise margin makes unimportant any error in the cut-in (enable) threshold, inherent in the simplified piecewise-linear model.

On setting $v_{BE} = e_{EO} - \Delta V_N$, we can find the maximum value of E_2 that will secure a reliable opening of the switch:

$$E_{2,\,max} = e_{EO} - \Delta V_N - I_{CBO,\,max} R_B$$

For germanium transistors, $E_{2,\,max}$ is always negative, because e_{EO} is small in comparison with I_{CBO}. For silicon transistors, $E_{2,\,max}$ is positive, provided ΔV_N is not very large.

The approximate calculation of a transistor switch outlined above is sufficiently accurate for most practical applications. If and where necessary, it can be made rigorous by using Eqs. (3.20), (3.24), and (3.26).

In more detail, the noise margin of transistor switches in cascade is discussed in Chap. 5.

Dynamic Models of Bipolar Transistors

The transit-time effects in a transistor become pronounced when the switch has to handle fast-changing signals and are due to the re-charging of C_E and C_C and the accumulation of minority carriers injected into the base. In some cases (such as in integrated-circuit transistors), an important role is also played by the minority carriers injected from base to collector.

Consider the behaviour of a switching transistor. Let the transistor be initially at cutoff. A switching (control) signal (such as a voltage step input) will cause the voltage across the barrier capacitances to change gradually. If the input signal has appropriate characteristics, this process will end in rendering one of the junctions conducting. To make our discussion more specific, let the transistor be forward-biased; then the emitter junction will be conducting. Since the voltage across a conducting p-n junction varies only slightly, C_e will have practically no effect on the further transient response of the circuit.

Minority-carrier injection through the forward-biased emitter junction leads to a gradual build-up of charge Q in the base and, as a consequence, a gradual increase in the reverse current through the reverse-biased collector junction, i_{qC}. The rise of current in the switched circuit is accompanied by a fall of collector-to-emitter voltage and a re-charging of the collector junction capacitance C_c. The process ends in forward-biasing the collector junction, which is another way of saying that the transistor moves into the saturation region.

At saturation, the voltage between the transistor terminals remains practically unchanged, so the junction capacitances have no effect on the transients. The base region keeps accumulating the minority-carrier charge Q which gradually approaches a level such that a state of dynamic equilibrium is attained between carrier injection, extraction, and recombination.

Thus, in turning on, a switching transistor traverses three regions: the cutoff region, the active region, and the saturation region. In turning off, a switching transistor goes through the same regions in the reverse sequence, with the saturation region coming first, the active region second, and the cutoff region last. It is to be noted that in a real transistor the transitions between regions at either

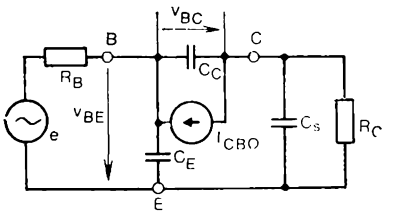


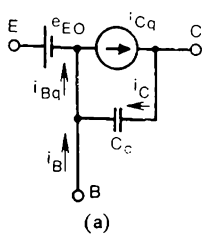
Fig. 3.17

turn-on or turn-off are more or less gradual, as depicted by the gradual variation in the slope of the *V-I* curves for the respective junctions. We shall consider dynamic models adapted to the piecewise-linear approximation of the transistor characteristics. In such models, the transitions between the various regions are abrupt, and the transient response associated with each transition is described by a linear differential equation.

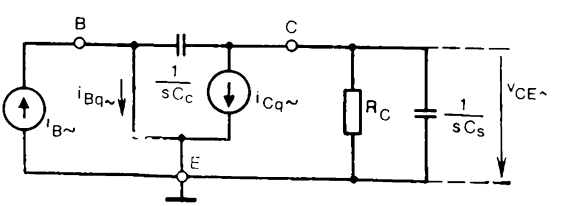
Given input signals, the dynamic model of a transistor for a particular region of operation (cutoff, active, or saturation) must, firstly, be able to bring out the manner in which the quantities characterizing the region in question vary, and, secondly, define the conditions required for the transistor to move to another region.

The cutoff region. Since no minority carriers are injected at cutoff, the inertia of the switching transistor is solely decided by the emitter and collector transition capacitances, C_e and C_c, so they must be included in the equivalent circuit (see Fig. 3.15*a*). Because the emitter and collector junctions, D_{E1} and D_{C1}, are reverse-biased (in the OFF state), the dynamic equivalent circuit of the transistor at cutoff can be constructed as shown in Fig. 3.17 which also includes the external circuit of a CE switch. From this equivalent circuit, we can readily deduce the manner in which v_{BE} and v_{BC} vary. For the transistor to move into the forward active region, it is required that $v_{BE} = e_{EO}$, and for the transistor to move into the reverse active region, it is required that $v_{BE} = e_{CO}$.

The active region. To make the situation more definite, we shall consider the forward active region. Since the emitter junction, D_{Ei}, is forward-biased, the generalized equivalent circuit of Fig. 3.15*a* can be re-arranged as shown in Fig. 3.18*a*. Under dynamic conditions variations in i_{qC} are associated with the storage of the minority-carrier charge Q in the base. For the emitter injection factor equal



(a)



(b)

Fig. 3.18

to $\gamma = 1$, variations in Q can quite accurately be described by the equation of charge

$$dQ/dt = -Q/\tau_\beta + i'_E - i_{qC}$$

where τ_β is the constant equal to the mean life time of minority carriers in the base. The three terms on the right-hand side of the equation define three factors affecting the variations of charge in the base region. The first term defines the rate of fall in the charge owing to recombination, the second refers to the rate of rise in the charge owing to carrier injection from emitter to base, and the third defines the rate of fall in the charge owing to carrier extraction from base to collector. The difference $i'_E - i'_{qC}$ gives the base current component i_{qB} (see Fig. 3.18*a*) which owes its existence to the majority-carrier charge stored in the base and balancing out the minority-carrier charge Q. Then, the equation of charge can be re-written as

$$dQ/dt + Q/\tau_\beta = i_{qB} \tag{3.29}$$

As is explained by transistor theory [14, 53], i_{qC} in a first approximation is proportional to Q. The relation connecting Q and i_{qC} can be formed from analysis of the steady state in which $i_{qC} = \beta i_{qB}$. On equating dQ/dt in Eq. (3.29) to zero (the condition of stationarity), we have

$$Q/\tau_\beta = i_{qB}$$

whence

$$Q = i_{qC}\tau_\beta/\beta \text{ or } i_{qC} = Q\beta/\tau_\beta \tag{3.30}$$

The relations defined by Eqs. (3.29) and (3.30) remain valid for $\gamma < 1$ as well, but τ_β becomes shorter than the mean life time of minority carriers in the base. On substituting the first equality of Eq. (3.30) into Eq. (3.29), we get

$$\tau_\beta (di_{qC}/dt) + i_{qC} = \beta i_{qB} \tag{3.31}$$

which directly connects the currents in the equivalent circuit of Fig. 3.18*a*.

It is convenient to analyze the transients in the circuit of Fig. 3.18*a* in terms of Laplace transforms. On expressing $i_{qC}(t)$ in Eq. (3.31) as

$$i_{qC\sim}(t) = i_{qC}(t) - i_{qC}(0)$$

and taking the Laplace transforms of the original time functions, we can re-write Eq. (3.31) as

$$i_{qC\sim}(s) = \beta(s)\, i_{qB\sim}(s) \tag{3.32}$$

where $\beta(s) = \beta/(1 + s\tau_\beta)$ is the Laplace transform of the CE current transfer ratio, and $i_{qB\sim}(s) = i_{qB}(s) - i_{qC}(0)/\beta$ is the Laplace transform of the base current increment relative to the initial steady-

state value. An equivalent transistor circuit utilizing the respective Laplace transforms appears in Fig. 3.18*b*. It illustrates the interrelation between $i_{qC\sim}$ and $i_{qB\sim}$, on the one hand, and $v_{CE\sim}$, on the other, in the amplification mode. It also includes R_C and $1/sC_s$ which simulate the circuit controlled by a CE transistor switch. The control circuit is represented by a current generator, $i_{B\sim}(s)$. Since the equivalent circuit has only been developed for incremental quantities, it does not contain d.c. supplies: E_C, I_{CBO}, and e_{EO}.

The condition for the transistor to move from the active into the saturation region is the same as that for forward-biasing the collector junction, that is

$$v_{CE}(t) = v_{CE(sat)} \approx 0$$

or, using the incremental collector-to-emitter voltage,

$$v_{CE\sim}(t) \approx -v_{CE}(0) \tag{3.33}$$

The condition for a transition from the active to the cutoff region is the same as that for reverse-biasing the emitter junction, which requires a negative emitter current:

$$|-i_E/i_{qC}| \approx 2\beta_I/\beta$$

At the instant when the transistor is reverse-biased, i_{qC} is connected to the turn-off base current by a relation of the form

$$i_{qC}(1 + 2\beta_I/\beta) = -i_{qB} \tag{3.34}$$

With drift transistors for which $\beta_I \ll \beta$, a transition to cutoff occurs at an emitter current practically equal to zero, so that $i_{qC} \approx -i_{qB}$.

The inequality $i_{qC} \neq 0$ is an indication that not all of the minority carriers stored have been swept out of the base region. Some of them remain in and are swept from the base region after the switch has been turned off, so a closer look must be taken at the transients during cutoff.

Dynamic cutoff. The transients existing during the final stage of turn-off (at dynamic cutoff) can conveniently be examined using the model shown in the equivalent circuit of Fig. 3.19. It differs

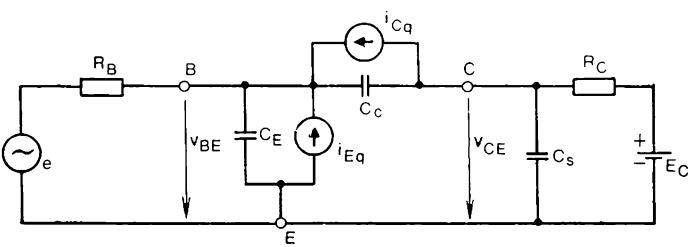


Fig. 3.19

from the equivalent circuit of Fig. 3.17 in that it includes generators of reverse collector and emitter currents, i_{qC} and i_{qE}, which decrease as the minority-carrier charge stored in the base decreases. To define the manner in which these currents change, we assume that they are related in the same fashion as at the instant of transistor turn-off, that is

$$i_{qE}/i_{qC} \approx 2\beta_I/\beta$$

so that the relation between i_{qB} and i_{qC} is represented by Eq. (3.34). Using Eq. (3.34), we can express the base current in Eq. (3.31) in terms of i_{qC}. On solving the resultant equation, subject to the initial condition, we get

$$i_{qC}(t) = | i_{qB}(0) | \frac{\beta}{\beta + 2\beta_I} \exp\left[-\frac{t}{\tau_\beta/(1+\beta)(1+2\beta_I)} \right] \qquad (3.35)$$

where $i_{qB}(0)$ is the turn-on base current at the instant when the transistor moves to dynamic cutoff.

The saturation region. Because the internal resistances of the forward-biased junctions are small, the currents i_C, i_E and i_B at saturation are practically independent of the internal state of the transistor, but solely decided by the external circuit parameters. The sole purpose of transient analysis in the saturation region is to determine the instant when the transistor moves out of saturation, so that the turn-off delay time can be found.

We shall use a very popular technique based on the use of Eq. (3.29) which describes variations in the minority-carrier charge in a transistor. A distinction of the saturation region is that the accumulation and removal of the saturation charge of minority carriers is accompanied by carrier injection across the emitter and collector junctions. Because of this, the charge density distribution in the material markedly differs from that existing in the active region, and the mean life time of minority carriers at saturation, τ_{sat}, is other than the time constant τ_β. Also, in the saturation region, there are no reverse currents flowing and no currents re-charging the junction capacitances so that $i_{qB} = i_B$. In view of the foregoing, the charge equation takes the form

$$dQ/dt + Q/\tau_{sat} = i_B \qquad (3.36)$$

For non-drift transistors, $\tau_{sat} \approx 0.5\tau_\beta$ to $0.9\tau_\beta$, whereas for drift transistors $\tau_{sat} \approx 1.5\tau_\beta$ to $2\tau_\beta$ or even more. The point is that because of the relatively high collector resistance the charge in such transistors is accumulated in both the base and the collector region.

Under steady-state conditions, that is, at $dQ/dt = 0$, we get

$$Q_{ss} = i_B\tau_{sat} \qquad (3.37)$$

Thus, the steady-state charge in a transistor is proportional to i_B and τ_{sat}.

At the boundary between the active and saturation regions, $i_B = I_{C(sat)}/\beta$, and the boundary charge is

$$Q_b = I_{C(sat)}\tau_{sat}/\beta \tag{3.38}$$

The transition from the amplification to the active region occurs at the instant when $Q(t) = Q_b$.

The uncertainties in the model represented by Eqs. (3.36) and (3.38) are mainly due to the fact that it does not allow for the nonlinearity of charge accumulation in real transistors—a factor which makes τ_{sat} a function of the operating conditions. Among other things, it ignores the marked change in the mean life time of carriers as the transistor moves from the active region to saturation. This is the reason why the boundary charge deduced from, say, Eq. (3.30) on substituting $i_{qC} = I_{C(sat)}$ differs from that yielded by Eq. (3.38). To avoid this discrepancy, let the total charge at saturation be the sum of boundary charge, $Q_b = I_{C(sat)}\tau_\beta/\beta$, and an excess charge, Q_{exc}, that is, $Q = Q_b + Q_{exc}$. Also, let the mean life time of the carriers constituting Q_b be equal to τ_β, because their concentration distribution is the same as in the active region, and the mean life time of the carriers constituting Q_{exc} be equal to τ_{sat}. Then the equation of charge can be re-cast as

$$dQ_b/dt + dQ_{exc}/dt + dQ_b/\tau_\beta + Q_{exc}/\tau_{sat} = i_B$$

Expressing Q_b in terms of $I_{C(sat)}$ gives

$$\tau_{sat}(dQ_{exc}/dt) + Q_{exc} = \tau_{sat}\left[i_B - \frac{1}{\beta}\left(I_{C(sat)} + \tau_\beta \frac{dI_{C(sat)}}{dt}\right)\right] \tag{3.39}$$

The condition for a transistor to move from saturation into the active region is

$$Q_{exc} = 0 \tag{3.40}$$

It can be shown that at a constant $I_{C(sat)}$, the refined model described by Eqs. (3.39) and (3.40) yields the same turn-off delay time as the model described by Eqs. (3.36) and (3.39).

Transient Response of a CE Transistor Switch

The response of a transistor switch (see Fig. 3.16) to a rectangular driving voltage $e(t)$, with allowance for the capacitive load component C_s is shown in Fig. 3.20. At time $t < t_0$, the driving voltage is $e = -E_2$, and the transistor is at cutoff. During the interval from t_0 to t_4, the switch is driven by a cut-in voltage, $e = E_1$, which causes the transistor to move into the saturation region (the switch

closes). At $t > t_4$, the driving voltage is again equal to $-E_2$ which causes the switch to open.

Let us take a closer look at what happens during the various switching times.

The delay time. This interval extends from t_0 to t_1, and its duration depends on the value of v_{BE}. Applying Kirchhoff's laws for the equivalent circuit of Fig. 3.17, we take the Laplace transform of the incremental voltage across the emitter junction, $v_{BE}(s)$, assuming that the Laplace transform of the change in the input signal is $(E_1 + E_2)/s$:

$$v_{BE}(s) = (E_1 + E_2)\frac{1 + sR_C(C_c + C_s)}{s(1 + as + bs^2)}$$

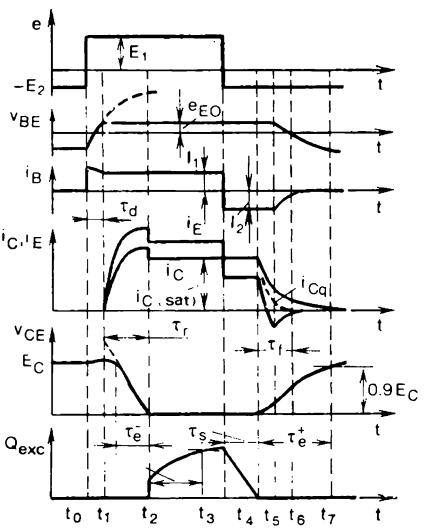


Fig. 3.20

where

$$a = R_B(C_e + C_c) + R_C(C_s + C_c)$$
$$b = R_B R_C(C_e C_c + C_e C_s + C_c C_s)$$

Time is counted from the beginning of the transient response in question.

Factoring the characteristic polynomial yields

$$v_{BE}(s) = (E_1 + E_2)\frac{1 + sR_C(C_c + C_s)}{s(1 + s\tau_1)(1 + s\tau_2)} \tag{3.41}$$

where $\tau_{1,2} = 0.5a\,(1 \pm \sqrt{1 - 4b/a^2})$. In practice, it usually happens that $R_B \gg R_C$, so that $4b/a^2 \ll 1$, and we may retain only the first two terms in the expansion of the radican into a power series, to get:

$$\tau_1 \approx a$$

and

$$\tau_2 \approx b/a \ll \tau_1$$

The original time function for the Laplace transform, Eq. (3.41), is

$$v_{BE\sim}(t) = (E_1 + E_2)[1 - A\exp(-t/\tau_1) - (1 - A)\exp(-t/\tau_2)]$$

where

$$A = [\tau_1 - R_C(C_c + C_s)]/(\tau_1 - \tau_2)$$

Recalling the condition for the transistor to turn on, the delay time τ_d can be found from the equation

$$v_{BE}(0) + v_{BE\sim}(\tau_d) = e_{EO}$$

where $v_{BE}(0)$ is determined by the initial steady state [see Eq. (3.28)]. Since $v_{BE\sim}(t)$ is a transcendental time function, this equation can readily be solved graphically or numerically. An approximate solution sufficiently accurate for practical purposes can be obtained, using the following approximations

$$\begin{aligned} 1 - A \exp(-t/\tau_1) &\approx 1 - A + At/\tau_1 \\ (1 - A) \exp(-t/\tau_2) &\approx (1 - A)/(1 + t/\tau_2) \end{aligned} \tag{3.42}$$

The original transcendental equation is replaced by a quadratic one, the positive root of which gives the sought result:

$$\tau_d \approx \frac{\lambda\tau_1 - \tau_2}{2A} + \sqrt{\frac{(\lambda\tau_1 - \tau_2)^2}{4A^2} + \frac{\lambda\tau_1\tau_2}{A}} \tag{3.43}$$

where

$$\lambda = \frac{e_{EO} - v_{BE}(0)}{E_1 + E_2} = \frac{E_2 + R_B I_{CBO} + e_{EO}}{E_1 + E_2}$$

To minimize τ_d, we must reduce λ, which can be done either by bringing down the cutoff voltage, E_2, or by raising the cut-in voltage, E_1.

The rise time (the interval from t_1 to t_2). After the emitter junction is rendered conducting, the base voltage remains practically constant ($v_{BE} = e_{EO}$), and the base current is

$$I_1 = (E_1 - e_{EO})/R_B$$

To evaluate the transient response of the transistor switch in terms of $v_{CE}(t)$ and $i_{qC}(t)$, we shall use the dynamic equivalent circuit shown in Fig. 3.18*b*. Since $i_{qC}(0) = 0$, we may write that

$$i_{qC\sim} = i_{qC} \quad \text{and} \quad i_{B\sim} = i_B$$

Using the equation of current balance at nodes B and C and setting $i_{B\sim}(s) = I_1/s$, we obtain an equation in the Laplace transform $v_{C\sim}(s)$:

$$v_{C\sim}(s)(1/R_C + sC_s + sC_c) = -\beta(s)[I_1/s - sC_c v_{C\sim}(s)]$$

Hence,

$$v_{C\sim}(s) = -(\beta I_1 R_C/s)[s^2\tau_\beta R_C(C_c + C_s) + s\tau_\beta^* + 1]^{-1}$$

where

$$\tau_\beta^* = \tau_\beta + R_C[C_c(1 + \beta) + C_s]$$

On evaluating the roots of the characteristic equation as before, we get

$$v_{C\sim}(s) = -\beta I_1 R_C[s(1 + s\tau_1^*)(1 + s\tau_2^*)]^{-1} \tag{3.44}$$

where

$$\tau_1^* \approx \tau_\beta^*, \quad \tau_2^* \approx (\tau_\beta/\tau_\beta^*)\, R_C\,(C_c + C_s)$$

such that $\tau_1 \gg \tau_2$.

Since τ_2 is negligibly small, the term $1/(1 + s\tau_2^*)$ in Eq. (3.44) can be replaced by the exponential function $\exp(-s\tau_2)$, approximately treating the time constant τ_2 as the delay of the exponential term with the time constant $\tau_1^* \approx \tau_\beta^*$:

$$v_{C\sim}(t) \approx -\beta I_1 R_C \left(1 - \exp \frac{t - \tau_2^*}{\tau_\beta^*}\right) \tag{3.45}$$

The collector current rise time, τ_r can be found from the condition defined by Eq. (3.33). On setting $v_{CE}(0) = E_C$, which corresponds to the initial (OFF) state of the switch, we get

$$-v_{CE\sim}(\tau_r) = E_C$$

the solution of which has the form

$$\tau_r = \tau_2 - \tau_\beta^* \ln(1 - E_C/\beta I_1 R_C) = \tau_2 - \tau_\beta^* \ln(1 - 1/K_{sat}) \tag{3.46}$$

The rise time decreases with increasing I_1. Under large-signal conditions ($K_{sat} \gg 1$), Eq. (3.46) may be replaced by an approximate equality

$$\tau_r \approx \tau_2^* + \tau_\beta^*/K_{sat} \tag{3.47}$$

As follows from the equivalent circuit of Fig. 3.18*b*, variations in i_{qC} are defined by a Laplace transform equation

$$i_{qC}(s) = -v_{C\sim}(s)\,[1 + sR_C\,(C_c + C_s)]/R_C$$

for which the original time function is

$$i_{qC}(t) = \beta I_1\,[1 - A_1 \exp(-t/\tau_\beta^*) - (1 - A_1)\exp(-t/\tau_2)] \tag{3.48}$$

where

$$A_1 = [\tau_\beta^* - R_C\,(C_s + C_c)]/(\tau_\beta^* - \tau_2^*) \tag{3.49}$$

At $t = \tau_r$ (the time t_2 on the waveform plots), the exponential term $\exp(-t/\tau_2^*)$ is negligibly small. Therefore, on substituting Eq. (3.47) into Eq. (3.48) and approximating $\exp(-\tau_r/\tau_\beta^*)$ by a linear function, $1 - \tau_r/\tau_\beta^*$, we may write i_{qC} at the instant when the transistor moves into the saturated region as

$$i_{qC}(\tau_r) \approx I_{C(sat)}\,[1 + (K_{sat} - 1)\,R_C\,(C_c + C_s)/\tau_\beta^*] \tag{3.50}$$

This current exceeds $I_{C(sat)}$ by an amount equal to the value of current flowing through C_c and C_s. The collector current is smaller than i_{qC} by an amount equal to the current flowing through C_c, that is

$$i_C(\tau_r) \approx I_{C(sat)}\,[1 + (K_{sat} - 1)\,R_C C_c/\tau_\beta^*].$$

At saturation, i_C abruptly changes to $I_{C(sat)}$ (see Fig. 3.20).

The interval from t_0 to t_2 may be treated as the turn-on time equal to

$$t_{ON} = \tau_d + \tau_r$$

Excess-charge accumulation time (the interval from t_2 to t_3). At time t_2, immediately after the transistor has moved into the saturated region, the charge of excess minority carriers in the base is

$$Q_{exc}(t_2) = Q(t_2) - Q_b = [i_{qC}(\tau_r) - I_{C(sat)}]\,\tau_\beta/\beta$$

Using Eq. (3.50) and expressing K_{sat} in terms of i_B and $I_{C(sat)}$, we find that

$$Q_{exc}(t_2) = \tau_2^*(I_1 - I_{C(sat)}/\beta)$$

According to Eq. (3.39), the steady-state value of the excess charge is

$$Q_{exc,\,ss} = \tau_{sat}(I_1 - I_{C(sat)}/\beta) \tag{3.51}$$

The accumulation of excess charge can be described by an exponential function of the form

$$Q_{exc}(t) = Q_{exc,\,ss} - [Q_{exc,\,ss} - Q(t_2)]\exp[-(t-t_2)/\tau_{sat}]$$

The accumulation interval is practically terminated at $Q_{exc} = 0.9Q_{exc,\,ss}$.

The switch remains in the steady ON state until the time t_3 when the command (switching) voltage is reversed in polarity, and a negative base cutoff current acts on the transistor. This current remains constant until the time t_5 when the emitter junction is reverse-biased:

$$i_B = -I_2 = -(E_2 + e_{EO})/R_B$$

The storage time (the interval from t_3 to t_4). During this interval, the saturation charge of excess minority carriers is removed from the base. According to Eq. (3.39), at $i_B = -I_2$, variations in the excess charge can be described by an exponential function of the form

$$Q_{exc}(t) = Q_{exc}(\infty) + [Q_{exc}(t_3) - Q_{exc}(\infty)]\exp(-t/\tau_{sat})$$

where $Q_{exc}(t_3) \approx Q_{exc,\,ss}$ is given by Eq. (3.51), and $Q_{exc}(\infty) = -\tau_{sat}(I_2 + I_{C(sat)}/\beta)$.

Using Eq. (3.40) and solving the resultant equation for the storage time $\tau_s = t_4 - t_3$, we obtain

$$\tau_s = -\tau_{sat}\ln\left[1 - \frac{I_{C(sat)}(K_{sat}-1)}{\beta(I_1+I_2)}\right] \tag{3.52a}$$

Equation (3.52a) is known as Moll's equation. Under large-signal conditions, when $I_{C(sat)}(K_{sat} - 1)/\beta(I_1 + I_2) \ll 1$, Moll's equation can be re-cast as

$$\tau_s \approx \tau_{sat} I_{C(sat)}(K_{sat}-1)/\beta(I_1+I_2) \tag{3.52b}$$

The storage time decreases with increasing I_2. Conversely, an increase in I_1 leads to an increase in τ_s.

The fall time (the interval from t_4 to t_6). In the general case, the fall time of i_{qC} is a sum of two intervals, namely from t_4 to t_5 (during which the transistor is in the active region) and from t_5 to t_6 (during which the transistor is at dynamic cutoff). The end of the fall time is defined by the instant when

$$i_{qC}(t_6) = 0.1 I_{C(sat)} \tag{3.53}$$

Let us take a closer look at the transient $i_{qC}(t)$ during the interval from t_4 to t_5 and find its duration, τ_{f1}. Since initially $i_{qC}(0) = I_{C(sat)}$ (the time t_4 is taken as the origin), the change in base current is

$$i_{B\sim}(t) = -(I_2 + I_{C(sat)}/\beta)$$

and variations in $i_{qC\sim}(t)$ are described by an expression analogous to Eq. (3.48)

$$i_{qC\sim}(t) = -(\beta I_2 + I_{C(sat)})[1 - A_1 \exp(-t/\tau_\beta^*) - (1 - A_1)\exp(-t/\tau_2^*)] \tag{3.54}$$

On setting $i_B(t) \approx i_{qB}(t)$ and using the condition for a transistor to move into the dynamic cutoff region, Eq. (3.34), we get

$$-i_{qC\sim}(\tau_{f1}) = I_{C(sat)} - I_2\beta/(\beta + 2\beta_I)$$

This equation defines τ_{f1} on the condition that its right-hand side is less than $0.9 I_{C(sat)}$. Otherwise, τ_{f1} will be bounded by Eq. (3.53). In the general case, τ_{f1} can be found from the following equation:

$$-i_{qC\sim}(\tau_{f1})/(\beta I_2 + I_{C(sat)}) = \lambda_1 \tag{3.55}$$

where

$$\lambda_1 = \begin{cases} \dfrac{I_{C(sat)} - I_2\beta/(\beta + 2\beta_I)}{\beta I_2 + I_{C(sat)}} & \text{for } \dfrac{I_2\beta}{\beta + 2\beta_I} > 0.1 I_{C(sat)} \\ \dfrac{0.9 I_{C(sat)}}{\beta I_2 + I_{C(sat)}} & \text{for } \dfrac{I_2\beta}{\beta + 2\beta_I} \leqslant 0.1 I_{C(sat)} \end{cases}$$

which can be solved graphically or numerically.

Consider an approximate solution based on the piecewise approximation of the function $i_{qC\sim}(t)$. To begin with, we replace the exponential terms of Eq. (3.54) by approximate expressions:

$$\exp(-t/\tau_\beta^*) \approx \begin{cases} 1 - t/\tau_\beta^* & \text{at } t < 2\tau_2^* \\ \exp(-t/\tau_\beta^*) & \text{at } t \geqslant 2\tau_2^* \end{cases}$$

$$\exp(-t/\tau_2^*) \approx \begin{cases} (1 - t/2\tau_2^*)/(1 + t/2\tau_2^*) & \text{at } t < 2\tau_2^* \\ 0 & \text{at } t \geqslant 2\tau_2^* \end{cases}$$

On either interval of approximation, an analytical solution of Eq. (3.55) can be obtained:

$$\tau_{f1} = \begin{cases} \tau_2^* - R_C(C_s + C_c) - \tau_\beta^* \ln(1 - \lambda_1) & \text{at} \quad \tau_{f1} > 2\tau_2^* \\ \dfrac{\lambda_1 \tau_\beta^* - 2R_C(C_s + C_c)}{A_1} + \Bigg[\left(\dfrac{\lambda_1 \tau_\beta^* - 2R_C(C_s + C_c)}{2A_1}\right)^2 & \\ + \dfrac{2\lambda_1 \tau_\beta R_C(C_c + C_s)}{A_1}\Bigg]^{1/2} & \text{at} \quad \tau_{f1} < 2\tau_2^* \end{cases} \tag{3.56}$$

In order to reduce τ_{f1}, we must reduce λ_1, which in turn requires that the ratio $I_2/I_{C(sat)}$ be increased.

During the interval from t_5 to t_6, the collector current falls in accordance with Eq. (3.35) as

$$i_C \approx i_{qC}(t) = I_2 \frac{\beta}{\beta + 2\beta_I} \exp\left(-\frac{t_6 - t_5}{\tau_{c\text{-}o}}\right)$$

where

$$\tau_{c\text{-}o} = \tau_\beta / (1 + \beta)(1 + 2\beta_I)$$

In view of Eq. (3.53), the duration of this interval is given by

$$\tau_{f2} = \begin{cases} \tau_{c\text{-}o} \ln \dfrac{10 I_2 \beta}{I_{C(sat)}(\beta + 2\beta_I)} & \text{at} \quad \dfrac{10 I_2 \beta}{I_{C(sat)}(\beta + 2\beta_I)} > 1 \\ 0 & \text{at} \quad \dfrac{10 I_2 \beta}{I_{C(sat)}(\beta + 2\beta_I)} < 1 \end{cases} \tag{3.57}$$

The total collector current fall time is

$$\tau_f = \tau_{f1} + \tau_{f2}$$

and the turn-off time of a transistor switch is

$$\tau_{OFF} = \tau_s + \tau_f$$

The large-signal case. In deriving Eqs. (3.56) and (3.57), we have assumed that λ_1 is positive, since otherwise τ_{f1} would be less than zero, which has no physical meaning in the case at hand.

It is seen from Eq. (3.55) that the condition $\lambda_1 > 0$ is equivalent to the restriction imposed on the turn-off current:

$$I_2 < I_{C(sat)}(1 + 2\beta_I/\beta)$$

When I_2 exceeds the above limit, the emitter junction is cut off before the collector junction is cut off. This form of transistor turn-off (called the inverted or emitter turn-off) is illustrated by the waveforms shown in Fig. 3.21.

Now, the transistor moves from saturation to the inverse active condition which is maintained over the interval from t_4 to t_5. At time t_5, the collector junction is reverse-biased, the transistor moves into the dynamic cutoff region, and the collector current decays exponentially with a time constant τ_{c-o}. The turn-off time for this

case can be found on assuming that the base current, $i_B \approx -I_2$, remains constant over the interval from t_4 to t_5, which is equivalent to satisfying the inequality

$$(\beta_I + 1)\, R_C \ll R_B$$

as is often the case in practice. In the circumstances, the interval from t_3 to t_5 (see Fig. 3.21) may be treated as the turn-off delay time, $\tau_{d,OFF} \approx \tau_s$, found from Moll's equations. The collector current fall time is

$$\tau_f \approx 2.3\tau_{c\text{-}o}$$

which corresponds to the decrease in i_{qC} to 10% of its initial value.

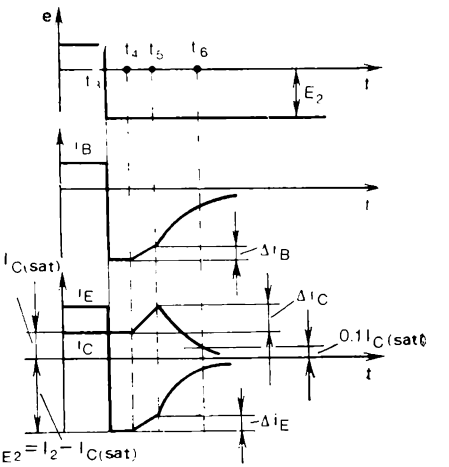


Fig. 3.21

The transition time (the interval from t_5 to t_7 in Fig. 3.20). During this interval, the transistor is in the dynamic cutoff region. As the collector current gradually decays to zero (I_{CBO} may be neglected) and the capacitances C_s, C_c and C_e are re-charged, the collector-to-emitter voltage, $v_{CE}(t)$, tends to a steady-state value, $v_{CE} = E_C$. The recovery stage is assumed to be completed when

$$v_{CE}(t_7) = 0.9E_C$$

The manner in which the collector voltage varies can be deduced from inspection of the equivalent circuit in Fig. 3.18*b*.

In practice, it often happens that $C_s \gg C_c$, $v_{CE}(t_5) \ll E_C$, and $R_C C_s \gg \tau_{c\text{-}o}$, so that it is safe to assume that over the interval from t_5 to t_6 the voltage varies exponentially with a time constant $R_C C_s$. Then,

$$\tau_t = R_C C_s \ln 10 \approx 2.3 R_C C_s \qquad (3.58)$$

The rise time, τ_r^+, of the transistor-switch output voltage v_{CE} is close to τ_t. As follows from Eq. (3.58), if we wish to cut down τ_t, we must reduce R_C. This, however, would entail an increase in power drain from the collector supply, E_C.

3.4. MODIFICATIONS OF TRANSISTOR SWITCHES

The basic CE transistor switch circuit can be modified in order to raise the speed of operation, mainly by cutting down τ_r, τ_s, and τ_f. Switch performance can be improved by optimizing the waveform

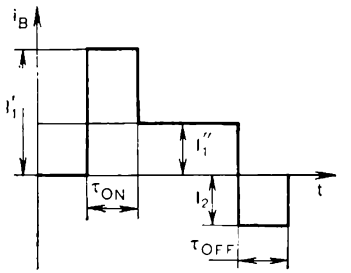


Fig. 3.22

of the base control current. An example of an optimum waveform is shown in Fig. 3.22. Because the rise time and the controlled part of the fall time of collector current decrease with an increase in the turn-on and turn-off currents, their absolute values, I_1' and I_2, must be high over the turn-on and turn-off intervals. On the other hand, a reduction in τ_s requires that K_{sat} be reduced too. For this to happen, the value of turn-on current, I_1', over the accumulation and steady-state ON intervals must be at a minimum still sufficient to saturate the transistor, that is,

$$I_1' = I_{C\,(sat)} K_{sat,\,min}/\beta_{min}$$

If the control voltage e is a rectangular waveform, the base current waveform can be made close to the optimal one, introducing a parallel RC-branch in the base circuit, and also by placing nonlinear elements (diodes) in the collector-base circuits.

Transistor switch with an RC control circuit (Fig. 3.23). This arrangement is employed in RC transistor logic (RCTL) circuits, flip-flops, monostable multivibrators, and similar circuits. C_B is a speed-up capacitor: it is used to permit faster turn-on of the transistor in response to a change in input and also helps to overcome the storage delay of the transistor itself. It may be added that when several such switches are connected in cascade, the principal component of C_s in the preceding stage is the C_B of the succeeding stage.

Let us examine the requirements that C_B must meet in order for the base current waveform to be close to optimal (see Fig. 3.23*b*).

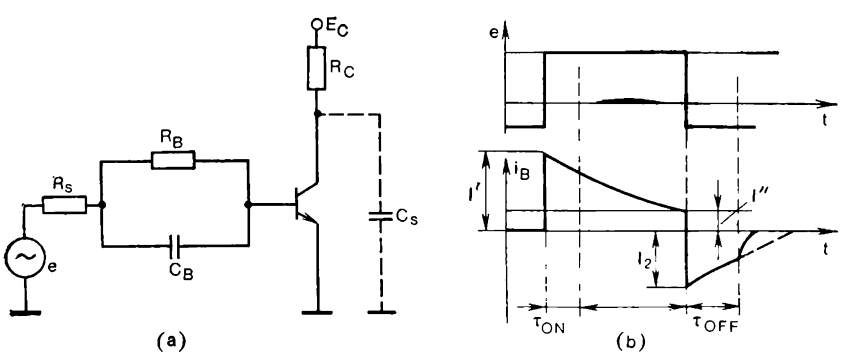


Fig. 3.23

The collector current and voltage rise time, τ_r, which accounts for the major part of the turn-on time, τ_{ON}, can be found by reference to the equivalent circuit of Fig. 3.18, recalling that the Laplace transform of base current is decided by the structure of the external base circuit components (see Fig. 3.23*a*):

$$i_B(s) = \frac{e(s)}{R'_s + R_B} \frac{1 + s C_B R_B}{1 + s C_B (R_B \parallel R'_s)} \tag{3.59}$$

where $R'_s = R_s + r'_B$.

The effect of C_B is especially strong during the initial interval which is small in comparison with the time constant $C_B (R_B \parallel R'_s)$. Thus, one condition for the choice of C_B is an inequality of the form

$$C_B (R_B \parallel R'_s) \gg \tau_{ON} \tag{3.60}$$

Then the base current transform is

$$i_B(s) \approx e(s)/R'_s = E_1/sR'_s \equiv I'_1/s$$

Obviously, the relations derived for a resistor-coupled switch will hold true in this case as well, on replacing I_1 by I'_1. The amount by which τ_r is thus cut down will depend on the ratio I'_1/I_1.

By time t_3, that is, the instant when a turn-off command is applied, the base current must fall to the value of the steady-state turn-on current:

$$I''_1 \approx E_1/(R_s + R_B) \approx E_1/R_B \quad (\text{at } R_B \gg R'_s)$$

Estimating the time it takes the charge on C_B to decay to the 5% level along an exponential curve, another condition for the choice of an optimal value for C_B is:

$$C_B (R_B \parallel R'_s) \leqslant \tau_p/3 \tag{3.61}$$

If the condition defined by Eq. (3.61) and also the condition for the maximum efficiency of the shaping circuit over the turn-off interval are both satisfied, we get

$$C_B (R_B \parallel R'_s) \gg \tau_{OFF} \tag{3.62}$$

The value of τ_s can be found from Moll's equations where

$$I_1 = I''_1$$
$$I_2 \approx [E_2 + v_{C_B}(\tau_p)]/R'_s$$

where $v_{C_B}(\tau_p) \approx E_1 R_B/(R'_s + R_B) \approx E_1$ is the voltage across the speed-up capacitor at the instant when the turn-off signal is applied, and $K_{sat} = \alpha I''_1/I_{C(sat)}$.

The duration of the controlled part of the collector-current trailing edge can be found from Eqs. (3.65) derived for the resistance-coupled switch, on substituting the above value of I_2.

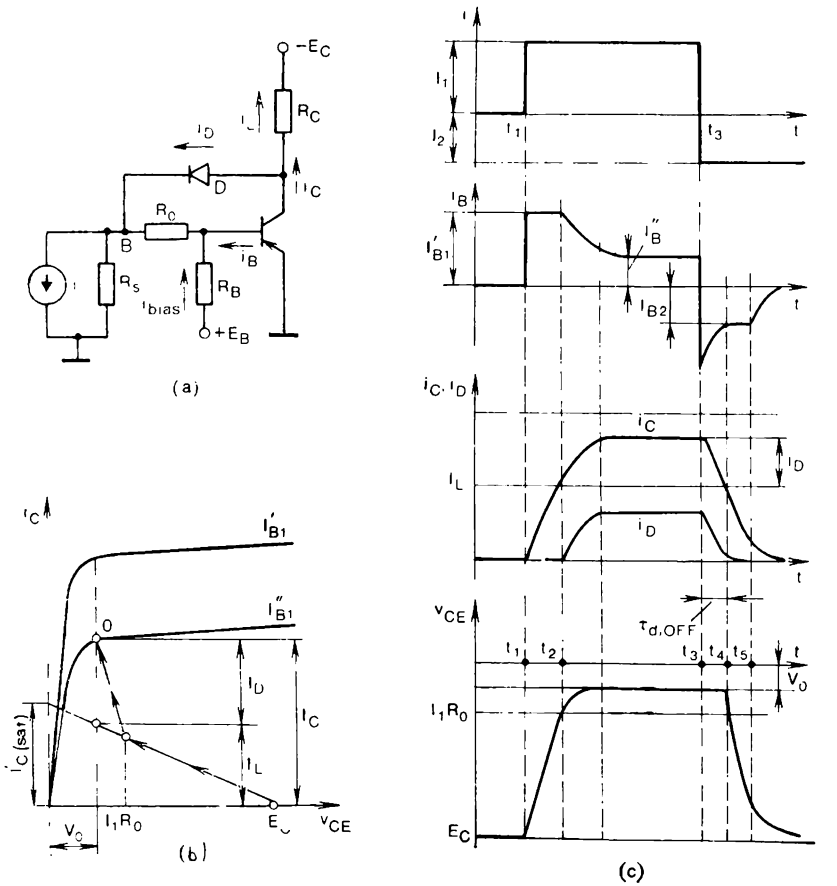


Fig. 3.24

Unsaturated switch with nonlinear negative feedback. In the transistor switches discussed so far, the storage time τ_s is a major or even, in some cases, the prevailing component of the overall switching time that limits the speed of the switch. To minimize the storage delay to a practically acceptable level, it is essential that in the ON state the transistor be in the active region. This calls for the stabilization of the collector potential so as to avoid any variations due to spread in parameters (mainly, β) and temperature variations. Placing a limiting diode [4] in the collector circuit would lead to a delay which is associated with the cutoff time of the diode when the switch is turned off, and might even exceed τ_s. To avoid this, the limiting diode is connected between the collector and base of the switching transistor (Fig. 3.24*a*).

When the switch is OFF, the diode is reverse-biased, so when a turn-on current step I_1 is applied at time t_1, the leading edge in accordance with the relations examined above for the control base current $I'_{B1} = I_1 - I_{bias}$, where $I_{bias} \approx (E_B + e_{EO})/R_B$. On the output characteristics (Fig. 3.24*b*), the operating point $(i_C(t), v_{CE}(t))$ moves along the load line $i_C = I_{C(sat)} - |v_{EC}/R_C|$ (if the load is purely ohmic). The diode is rendered conducting at the end of the leading edge of the collector voltage (time t_2 in Fig. 3.24*c*), where v_{CE} rises to $v_{CE}(t_2) \approx -I_1 R_0$. From that instant on, the input current is distributed with a time constant τ_β/β between the base and diode circuits so that the steady state is attained when the operating point is at 0 on the plot of Fig. 3.24*b*, where $I_C = I_L + I_B$ and $v_{CE} = V_0 \approx (I_{bias} + I'_{B1}) R_0 - v_D + e_{EO}$, where v_D is the voltage drop across the diode. With the operating conditions of the circuit adequately chosen, the point 0 will lie within the active portion of the characteristic, in close proximity to the saturation boundary, $V_{CE(sat)}$.

Because the base circuit has a relatively low resistance, arrival of a turn-off current step, $-I_2$, at time t_3 produces a base current step, $i_B(t_3) \approx -(I_2 + I_D + I_{bias})$, which marks the instant at which the collector current begins to fall. The delay in the collector voltage fall (which begins at time t_4) arises because it takes some time for the diode to turn off and for the collector current to fall by I_D.

The collector current and voltage fall at $t > t_5$ in the same manner as before, with the turn-off base current $I_{B2} = -(I_2 + I_{bias})$. Thus, an unsaturated switch with nonlinear negative feedback shows a certain turn-off delay, $\tau_{d,OFF}$, but it is substantially smaller than the storage time, τ_s, observed in saturated transistor switches.

An unsaturated switch with nonlinear negative feedback is roughly equivalent to a saturated switch with a speed-up capacitor (provided the speed-up capacitor has an optimal value), which is corroborated by the base current waveform in Fig. 3.24*c*, but it does not contain an extra capacitor.

Unsaturated Schottky transistor switch. In integrated circuits, nonlinear feedback is most efficient if the feedback loop contains a Schottky-barrier diode fabricated in the same chip with a Schottky transistor. As is seen from Fig. 3.25*a*, this arrangement eliminates R_0. This is achieved because the V/I characteristic of a Schottky-barrier diode lies to the left of the characteristic of the semiconductor diode formed by the collector p-n junction of the transistor (Fig. 3.25*b*). As a result, when a moderate current, I_D, flows through the Schottky-barrier diode, the operating point of the diode formed by the p-n junction (the dashed line in Fig. 3.25*a*) lies below the threshold voltage, so that in the ON state the switching transistor is in the active region close to the saturation boundary. Schottky transistor switches have a very short storage time, $\tau_s < 1$ ns. Also, they have a

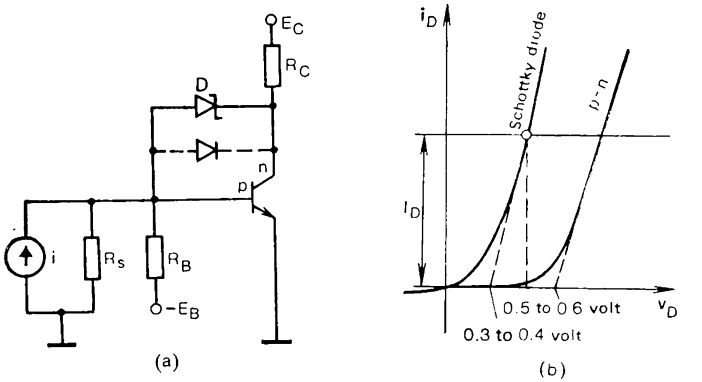


Fig. 3.25

higher β and lower uncontrolled currents, and a good temperature stability of their dynamic and static characteristics.

REVIEW QUESTIONS AND PROBLEMS

3.1. Point on the *V/I* characteristic of a diode the regions corresponding to the closed and open states of the diode switch.

3.2. Given a circuit consisting of a series combination of a voltage generator e, diode D and resistor R, determine the current through and the voltage across the diode and resistor for $e = \pm 10$ V, assuming that $R = 1$ kΩ and that the diode is a silicon type. Use the piecewise-linear diode model in the calculations.

3.3. How will the results of problem 3.2 change, if use is made of the ideal diode model, the exponential diode model?

3.4. Using the exponential diode model for problem 3.2, find the value of e corresponding to the switching point of the switching diode. Name the two likely methods for locating the switching point and compare the results.

3.5. Plot the transfer function for the series diode switch of Fig. 3.11*a*, which relates the voltage across R_L to the driving voltage. Assume that the diode is a germanium type, $R_{drive} = 20\ \Omega$, $R_{bias} = 2$ kΩ, $R_L = 10$ kΩ, and $E_{bias} = 5$ V. Analyze the circuit, using the piecewise-linear diode model. How does the resultant transfer function differ from that of a real series diode switch?

3.6. Plot the transfer function for the parallel diode switch of Fig. 3.11*b*, using a silicon diode and assuming that $R_{drive} = 1$ kΩ, $R_{bias} = 2$ kΩ, $R_L = 5$ kΩ, and $E_{bias} = -5$ V. Analyze the circuit, using the piecewise-linear diode model. How does the resultant transfer function differ from that of a real parallel diode switch?

3.7. Using ALGOL, FORTRAN or PL/1, present the piecewise-linear and exponential static models of the diode as a routine for calculating the diode current, taking the diode voltage as parameter.

3.8. Can the simplified piecewise-linear or ideal diode model be represented as a computer routine for calculating the diode current from a given voltage?

3.9. List the factors governing the speed of a diode switch during turn-on and turn-off. Name the intervals that make up the overall switching time of a diode switch.

3.10. Determine the mean life time of minority carriers in the base of an alloyed pulse diode for which the manufacturer's data-sheet quotes a reverse recovery time of 100 ns at a forward current of 20 mA, assuming that the recovery time is defined as the interval required for the reverse current of the diode to fall to 10% of its original value upon turn-off.

3.11. The diode switch of Fig. 3.12*a* uses a charge-storage diode for which $\tau = 1$ μs and $\gamma_E = 7$. Determine the diode voltage waveform, if e varies from +5 V to −5 V linearly during $\tau_r = 3$ μs. In calculation, neglect the effect of C_b and C_s, and also assume that R is such that $r_f \ll R \ll r_r$.

3.12. Determine the voltage waveform across the diode in the switch of the previous problem, if the input voltage is a sine wave with an amplitude of 10 V and a frequency of 0.5 MHz.

3.13. Using the generalized equivalent circuit for the piecewise-linear static model of a transistor, develop linear equivalent circuits for the cutoff, forward active, reverse active, and saturated conditions of operation.

3.14. Solve problem 3.13 for the simplified piecewise-linear transistor model.

3.15. Using the piecewise-linear transistor model, determine the incremental (dynamic) resistances between the collector and emitter and between the base and emitter at saturation.

3.16. Develop an equivalent circuit for a transistor in the forward active region, answering Eqs. (3.21) and (3.22), at $K_{sat} \approx 1$ and $v_{CE} \gg v_{BE}$. Represent the input circuit as a voltage generator v_{BE}.

3.17. Compare the equivalent circuits in Problems 3.14 and 3.16. How do they differ? Can the linear equivalent circuit be reduced to that in Problem 3.16 by equivalent transformation?

3.18. The switch in Fig. 3.15 has the following parameters: $R_B = 10$ kΩ, $R_C = 1$ kΩ, $E_C = 10$ V, $\beta = 50$, $\beta_I = 1$, $m\varphi_T = 30$ mV, $V_{BE\,(sat)} = 0.7$ V at $K_{sat} = 20$ and $I_{C\,(sat)} = 1$ mA. Determine v_{BE}, and v_{CE} at $e = 0$, 0.5, 1 and 5 V, neglecting the bulk resistances r_B, r_C, and r_E and the reverse resistances $r_{C,r}$ and $r_{E,r}$.

3.19. Determine the collector bulk resistance for a transistor with $\beta = 60$, $\beta_I = 2$, and $m\varphi_T = 30$ mV, if it is known that at I_B

$= 1$ mA and $I_C = 5$ mA, the collector-to-emitter voltage is $v_{CE} = 0.1$ V.

3.20. Given a transistor for which, at a base current of 1 mA, the absolute value of collector-to-emitter voltage with the collector circuit open is 40 mV and with the emitter circuit open, 1 mV. Using the exponential transistor model, find V_{CE} at $I_B = I_C = 2$ mA if $I_{CEO} \ll I_C$ and $m\varphi_T = 30$ mV.

3.21. Using ALGOL, FORTRAN or PL/1, present the piecewise-linear and exponential static models of a transistor as a computer routine for calculating i_C and i_E from the given v_{CB} and v_{EB}.

3.22. In the transistor switch of Fig. 3.16, $R_B' = 25$ kΩ, $E_C = 10$ V. The collector and emitter junction capacitances are given by $C_e = 14\,(V_{EBO} + 0.3)^{-1/2}$ pF and $C_c = 20\,(V_{CBO} + 0.3)^{-1/2}$ pF. The driving voltage e changes stepwise from -5 V to $+25$ V. Determine the total charge that must be injected into the base in order to bring the transistor to the cut-in boundary, and also the cut-in delay, neglecting the variations in the current over the delay interval. Determine the delay again, assuming the transition capacitances constant and equal to their maximum values. Assuming $R_C \ll R_B$, the effect of R_C may be neglected.

3.23. In the switch of Fig. 3.16 for which $R_B = 10$ kΩ, $R_C = 1$ kΩ, and $E_C = 12$ V, the transistor is initially at the cut-in boundary. It is known that a transistor for which $\beta = 30$ and $\bar{C}_c = 10$ pF enters the saturated region 0.3 μs after a step voltage of 10 V is applied to the switch input. Determine τ_β and f_{max} of the transistor, assuming that its static characteristics are represented by the ideal transistor model with sufficient accuracy.

3.24. An *n-p-n* transistor for which $f_{max} = 50$ MHz, $\beta = 40$, and $\bar{C}_c = 3$ pF, is operating in a switching circuit for which $E_C = 12$ V, $R_C = 1$ kΩ, and $C_s = 5$ pF. Determine the turn-on current, assuming that the duration of the negative leading edge of the collector voltage is 1 μs, 100 ns, and 10 ns.

3.25. Initially, the driving voltage applied to the input of the switch in Fig. 3.16 is $e = -4$ V. At time t_0, the input is a positive rectangular pulse during which $e = +6$ V. The pulse duration is 10 μs. Determine τ_d, τ_r, τ_s, τ_f, and τ_t, if $R_B = 2$ kΩ, $R_C = 250$ Ω, $E_C = 5$ V, the transistor is a germanium type for which $\beta = 60$, $f_{max} = 100$ MHz, $\bar{C}_e = 9$ pF, $\bar{C}_c = 6$ pF, $\tau_{sat} = 20$ ns, and $e_{EO} = 0.7$ V.

3.26. In the transistor switch of Fig. 3.23*a*, $E_C = 20$ V, $R_C = 2$ kΩ, $\beta = 60$, and $\tau_\beta = 15$ μs. The input signal is a voltage step of 5 V (the initial voltage being zero). Neglecting the transition capacitances and the internal resistance of the signal source, select R_B and C_B such that the collector current step is 5 mA and the collector cur-

rent following the step remains constant. In solving the problem, use the ideal transistor model.

3.27. In the transistor switch of Fig. 3.23*a*, $E_C = 12$ V, $R_C = 0.5\,\text{k}\Omega$, $R_B = 5\,\text{k}\Omega$, $C_B = 200$ pF, $\tau_\beta = 1\ \mu\text{s}$, $\tau_{sat} = 0.3\ \mu\text{s}$, $\beta = 50$, and the total resistance of the signal source and the base bulk resistance, r'_B, is 1 kΩ. The input voltage level corresponding to the OFF state is −1 V, and to the ON state, +5 V. The driving input pulse is a positive rectangular waveform with a duration of 20 μs. Neglecting the transition capacitances and the input resistance of the transistor, plot the base-current waveform, and determine the turn-on, accumulation and storage times.

4 LIMITING AND CLAMPING CIRCUITS

4.1. GENERAL

We shall begin this chapter by considering simple pulse shaping circuits which utilize the nonlinear behaviour of switching elements without positive feedback. This sets them apart from regenerative (positive-feedback) pulse-shaping circuits, such as flip-flops, multivibrators and blocking oscillators.

The basic type of a nonlinear switch is the *peak limiter**. It delivers an output waveform similar to the input waveform so long as the instantaneous value of the latter lies within the specified thresholds. Should the input signal exceed such a threshold, the output signal will be constrained to a constant value known as the limiting level.

The transfer characteristic of an ideal peak limiter consists of two straight-line segments, one sloping and the other horizontal. Limiters having the transfer characteristics shown in Fig. 4.1*a* and *b* are known as *positive* and *negative* limiters, respectively. Limiters having the transfer characteristic shown in Fig. 4.1*c* and *d* are known

* As often, it is called a clipper.—Translator's note.

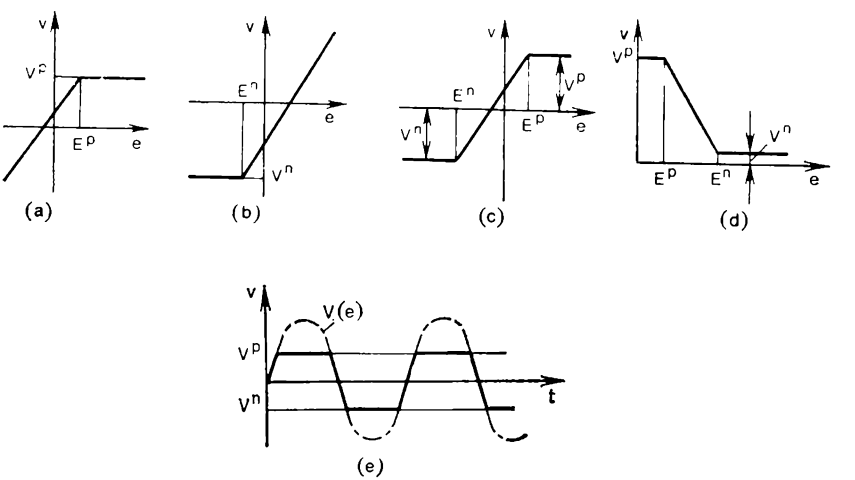


Fig. 4.1

as *double-peak* or positive-negative limiters. Figure 4.1*e* shows the output signal of a positive-negative limiter driven by a sinusoidal input.

Suppose that within the limiting thresholds (the transmission band), the output voltage of a limiter v is a linear function of the input voltage e:

$$v = V(e) = V_0 + Ke \tag{4.1}$$

Since the transfer characteristics of real limiters may have a finite slope within the limiting region, we also introduce linear functions

$$V^{\mathrm{p}}(e) = K^{\mathrm{p}}e + V_0^{\mathrm{p}}$$

and

$$V^{\mathrm{n}}(e) = K^{\mathrm{n}}e + V_0^{\mathrm{n}}$$

which approximate the relation $v(e)$ in the positive and negative limiting regions, respectively.

Assuming an arbitrary e, the transfer characteristic of a limiter can be written:

—for positive limiting

$$v = \min [V(e),\ V^{\mathrm{p}}(e)]$$

—for negative limiting

$$v = \max [V(e),\ V^{\mathrm{n}}(e)]$$

—for both positive and negative limiting

$$v = \min \{\max [V(e),\ V^{\mathrm{n}}(e)], V^{\mathrm{p}}(e)\}$$

For a limiter to operate normally, it is essential that

$$|K^{\mathrm{p}}/K| \ll 1,\ |K^{\mathrm{n}}/K| \ll 1, \quad V_0^{\mathrm{p}} > V_0^{\mathrm{n}} \tag{4.2}$$

The limiting thresholds E^{p} and E^{n} can be found from the equations

$$V(E^{\mathrm{p}}) = V^{\mathrm{p}}(E^{\mathrm{p}})$$

and

$$V(E^{\mathrm{n}}) = V^{\mathrm{n}}(E^{\mathrm{n}})$$

Hence,

$$\begin{aligned} E^{\mathrm{p}} &= (V_0^{\mathrm{p}} - V_0)/(K - K^{\mathrm{p}}) \\ E^{\mathrm{n}} &= (V_0^{\mathrm{n}} - V_0)/(K - K^{\mathrm{n}}) \end{aligned} \tag{4.3}$$

Because of stray reactances (usually, capacitances) and the finite switching times, the output signal reaches the limiting level(s) only after all transients have died out. In evaluating the speed of a limiter, it is customary to consider its transient response to an ideal rectangular input pulse.

4.2. DIODE LIMITERS

In a limiter circuit, the diode can be connected in parallel, series, or series-parallel with the load. Also, it may or may not contain an operational amplifier.

In a parallel positive limiter (Fig. 4.2*a*), the diode remains at cutoff, and the input voltage is allowed to pass on to the output so long as it is below what is often called the reference voltage, E. At $e > E$, the diode is driven into conduction, and the output voltage is held at $V^{p} = E$, that is, E is both the limiting threshold and the limiting level.

In a parallel negative limiter (Fig. 4.2*b*), the diode is driven to cutoff when $e < E$, so that the reference voltage E defines both the limiting threshold and the limiting level.

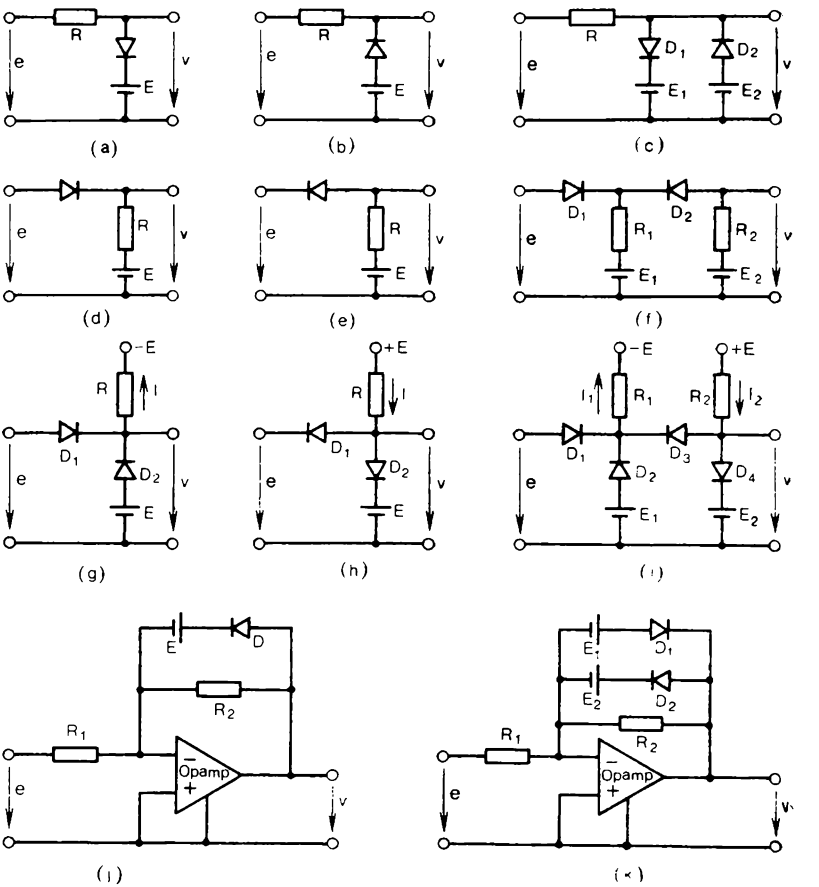


Fig. 4.2

In a positive-negative limiter (Fig. 4.2*c*), D_1 flattens the positive peaks at E_1, and D_2 flattens the negative peaks at E_2 ($E_1 > E_2$). When the inequality $E_1 > e > E_2$ is satisfied, both diodes are at cutoff, and the input voltage is passed onto the output.

Parallel limiters have a low (zero) output resistance when they limit the signal, and a high output resistance when they pass the signal.

Series limiters (Fig. 4.2*d* through *f*) pass the input signal when the diodes are conducting and limit the input signal when the diodes are at cutoff. The circuit of Fig. 4.2*d* performs positive limiting (with $e > E$) to $V^p = E$, whereas the circuit of Fig. 4.2*e* performs negative limiting (with $e < E$) to the same level. In the double-peak limiter of Fig. 4.2*f*, the right-hand diode, D_2, flattens the positive peak to a level E_1, and the left-hand diode, D_1, flattens the negative peak to a voltage level produced at the output when this diode is driven to cutoff, that is,

$$V^n = (E_1R_2 + E_2R_1)/(R_1 + R_2)$$

A series diode limiter has a low output resistance in passing the signal, and a high output resistance in limiting the signal.

Series-parallel limiters (Fig. 4.2*g* through *i*) are employed where it is desired to retain a low output resistance when both passing and limiting the input signal. For its operation, such a limiter depends on switching the current I maintained by a supply source E and a resistor R between the diodes D_1 and D_2 (see Fig. 4.2*g* and *h*). In passing the signal, the series diode D_1 is conducting, and in limiting the signal, the parallel diode D_2 is conducting. The double-peak limiter of Fig. 4.2*i* is a cascade combination of two simple limiters. The reference voltage E_2 sets the positive-limiting level, and E_1 sets the negative-limiting level. For signal limiting, the diodes D_2 and D_3 must be conducting at the same time, which can be obtained when $I_1 < I_2$.

In passing the input signal, the above limiters have a gain of $K \leqslant 1$. If it is desired that the signal is amplified within the passing region, the limiter must be extended to include an amplifier. The best arrangement is to place a switching element in the feedback path of an operational amplifier (Fig. 4.2*j* and *k*). With the switching element open, $K \approx -R_2/R_1$ [see Eq. (2.31)]. When the switch closes, the resistance in the feedback path falls to zero, and the output voltage remains unchanged. In the circuit of Fig. 4.2*j*, the diode turns on at $v < E$, which is equivalent to $e > -ER_1/R_2$. In other words, this circuit is a positive limiter with a limiting threshold $-ER_1/R_2$ and a limiting level of E. To obtain negative limiting, the diode must be connected in reverse polarity. Positive and negative limiting can be combined by placing two diodes in the feedback circuit. D_1 provides positive limiting, and D_2 negative limiting. Limiters

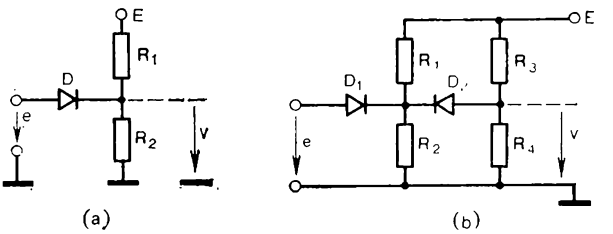


Fig. 4.3

containing operational amplifiers have a low output resistance within any portion of the transfer characteristic.

In series limiters in which the reference voltage source is returned to a common junction with R, the desired limit levels can conveniently be obtained by using resistive voltage dividers (Fig. 4.3). The circuit of Fig. 4.3*a* is a negative limiter to a level $V^n = ER_2/(R_1 + R_2)$, and the circuit of Fig. 4.3*b* is a double-peak limiter with a positive limit level $E^p = ER_4/(R_3 + R_4)$, and a negative limit level $E^n = E\,(R_2 \parallel R_4)/[(R_1 \parallel R_3) + (R_2 \parallel R_4)]$.

Static Characteristics of Diode Limiters

From an analysis of the static characteristic of a diode limiter, we can evaluate the effect of the residual parameters of the diode, load and the internal resistance of the signal source on the accuracy of limiter operation usually defined in terms of the stability of the transfer characteristic. We shall discuss the analysis procedure, taking a parallel positive limiter as an example.

A schematic diagram of such a limiter, extended to include the load resistance R_L and the internal resistance R_s of the signal source, appears in Fig. 4.4*a*. Using the piecewise-linear diode model shown in Fig. 3.6*b*, we can develop linear equivalent circuits for this limiter when it passes the applied signal (Fig. 4.4*b*) and when it limits

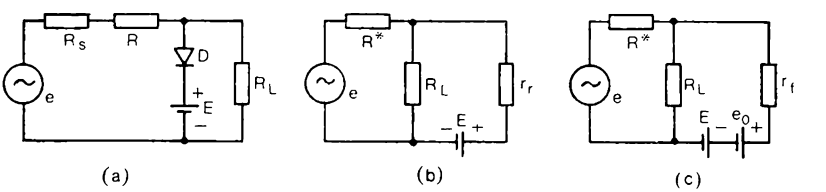


Fig. 4.4

the signal (Fig. 4.4*c*); in these equivalent circuits, $R^* = R_s + R$. The equivalent resistance of the circuit switched by the diode is determined by the parallel connection of R^* and R_L. Denoting $\gamma_f = r_f/(R^* \parallel R_L)$ and $\gamma_r = (R^* \parallel R_L)/r_r$, and referring to the equivalent circuits of Fig. 4.4*b* and *c*, we get

$$\begin{aligned} K &= \frac{R_L}{R^*+R_L}\,\frac{1}{1+\gamma_r}, & V_0 &= E\gamma_r/(1+\gamma_r) \\ K^p &= \frac{R_L}{R^*+R_L}\,\frac{\gamma_f}{1+\gamma_f}, & V_0^p &= (E+e_0)/(1+\gamma_f) \end{aligned} \tag{4.4}$$

In designing a diode limiter, the values of R and R_L are chosen such that

$$\gamma_f \ll 1, \quad \gamma_r \ll 1 \tag{4.5}$$

The first inequality ensures the correct operation of the limiter ($K^p/K \ll 1$), and the second eliminates the effect of the instability in the reverse resistance of the diode on the transfer characteristic of the limiter.

In view of Eqs. (4.5), Eqs. (4.4) can be re-cast in a simpler form

$$\begin{aligned} K &\approx R_L/(R^*+R_L), & V_0 &\approx E\gamma_r \\ K^p &= K\gamma_f, & V_0^p &\approx E+e_0 \end{aligned} \tag{4.6}$$

Using Eq. (4.2), we find the limiting threshold to be

$$E^p = [E(1-\gamma_r) - e_0]/K(1-\gamma_f)$$

The above relations show that, when the conditions defined by Eqs. (4.5) are satisfied, the principal destabilizing factor is the drift in the initial forward voltage of the diode, e_0.

Dynamic Characteristics of Diode Limiters

Since diode limiters use high-speed pulse diodes, the principal factor controlling their dynamic characteristics is the load capacitance in circuits without an operational amplifier, and the lag time associated with the amplifier in circuits using an operational amplifier. We shall carry out a comparative analysis of dynamic characteristics for various types of diode limiters without an operational amplifier, using as an example a positive limiter and assuming that it is driven by an ideal rectangular pulse in positive polarity.

Parallel limiter. Variations in the output voltage of a parallel limiter take place when the diode is in the OFF state. Applying Thevenin's theorem to the circuit of Fig. 4.4*b* and including the load capacitance C_L, the limiter may be represented by the equivalent circuit of Fig. 4.5*a* where

$$R_{eq} = R^* \parallel R_L \parallel r_r \approx R^*R_L/(R^*+R_L) = R^*K$$

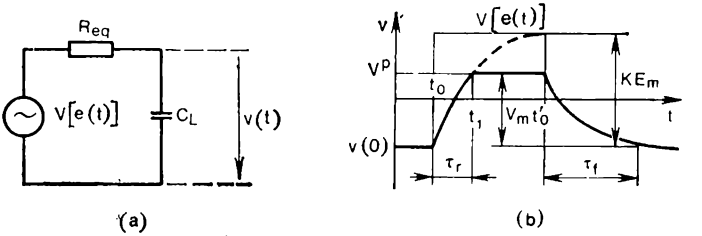


Fig. 4.5

The value of $V(e)$ is defined by Eqs. (4.1) and (4.4). When an input pulse is applied at time t_0, the output voltage begins to vary exponentially with a time constant $\tau_{eq} = R_{eq}C_L$:

$$v(t) = v(0) + KE_m\{1 - \exp[-(t - t_0)/\tau_{eq}]\} \tag{4.7}$$

where $v(0)$ is the initial steady-state value of output voltage, and E_m is the peak value of the input pulse.

Exponential variations in $v(t)$ take place over the interval from t_0 to t_1 (Fig. 4.5*b*) and finally cause the diode to conduct and the limiter to switch over to the limiting action. The rise time τ_r of the output voltage can be found from the equation

$$v_{\sim}(t_0 + \tau_r) = V_m$$

where V_m is the peak value of output voltage equal to $V^p - v(0)$. Solving this equation for the rise time yields

$$\tau_r = -\tau_{eq} \ln(1 - V_m/KE_m) \tag{4.8}$$

When the driving pulse ceases at time t_0', the output voltage begins to approach $V[e(t)] = v(0)$ asymptotically. Therefore, the fall time τ_f of the output pulse is independent of E_m and is equal to

$$2\tau_{eq} - 3\tau_{eq}.$$

Since for the circuit to operate as a limiter it is required that $V_m/KE_m < 1$, in a parallel limiter $\tau_r < \tau_f$ always.

Series limiter. In this type of limiter, the output voltage rises when the diode is in the OFF state, and falls when the diode is in the ON state. Applying Thevenin's theorem and including both the load and internal source resistances, we obtain the equivalent circuits shown in Fig. 4.6. The circuit in Fig. 4.6*a* applies when the output voltage is falling, and the circuit in Fig. 4.6*b* does when the output voltage is rising. The respective equivalent resistances, $R_{eq,\,ON}$ and $R_{eq,\,OFF}$, are given by:

$$R_{eq,\,ON} = (R_s + r_f) \parallel R \parallel R_L$$

$$R_{eq,\,OFF} = (R_s + r_r) \parallel R \parallel R_L \approx RR_L/(R + R_L)$$

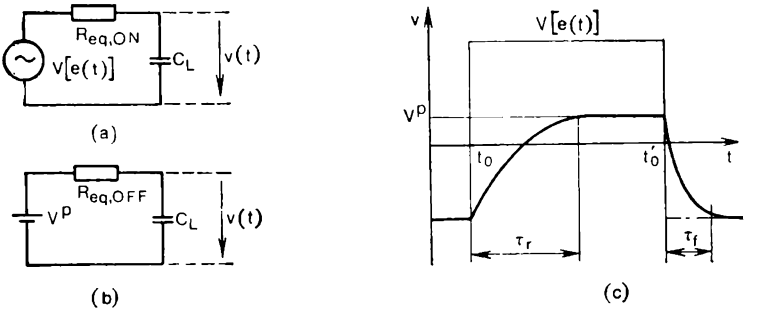


Fig. 4.6

In response to an input voltage step, the output voltage rises exponentially with a time constant $\tau_{OFF} = R_{eq,\,OFF}\,C_L$, and falls with a time constant $\tau_{ON} = R_{eq,\,ON}C_L$. In either case, $v\,(t)$ asymptotically approaches the respective steady-state value, V^p or $V\,[e(t)]$ at $t > t_0'$ (Fig. 4.6*c*). The rise and fall times of the output voltage are $\tau_r = 2\tau_{OFF}\text{-}3\tau_{OFF}$, and $\tau_f = 2\tau_{ON}\text{-}3\tau_{ON}$. Since $R_{eq,\,ON} < R_{eq,\,OFF}$, a series limiter differs from a parallel limiter in that $\tau_r > \tau_f$.

Series-parallel limiter. This configuration combines the properties of a parallel and a series circuit. The output voltage can vary only when the diode D_2 is OFF, whereas the diode D_1 is conducting during the fall in the output voltage and non-conducting during its rise. The equivalent circuit of the limiter during the fall interval is shown in Fig. 4.7*a*. The equivalent resistance $R_{eq,\,ON}$, combining the load and internal source resistances, is given by

$$R_{eq,ON} = (R_s + r_f) \,||\, R \,||\, R_L \,||\, r_r \approx (R_s + r_f) \,||R \,||\, R_L$$

During the fall interval ($t > t_0'$), the output voltage decreases exponentially (Fig. 4.7*c*) with a time constant $\tau_{ON} = R_{eq,\,ON}C_L$ and approaches asymptotically its steady-state value $V\,[e\,(t)]$. As in the series circuit, $\tau_f = 2\tau_{ON}\text{-}3\tau_{ON}$.

The equivalent circuit of Fig. 4.7*b*, applying to the rise interval, has been derived on the assumption that the reverse resistance r_r of the diodes at cutoff is negligibly small. Then,

$$R_{eq,\,OFF} = R_L \,||\, R \text{ and } E_{eq} = ER_{eq,OFF}/R$$

Over the rise interval ($t_0 < t < t_1$), the output voltage rises exponentially

$$v\,(t) = [v\,(0) - E_{eq}] \exp\left[-\frac{t - t_0}{C_L R_{eq,\,OFF}}\right] + E_{eq}$$

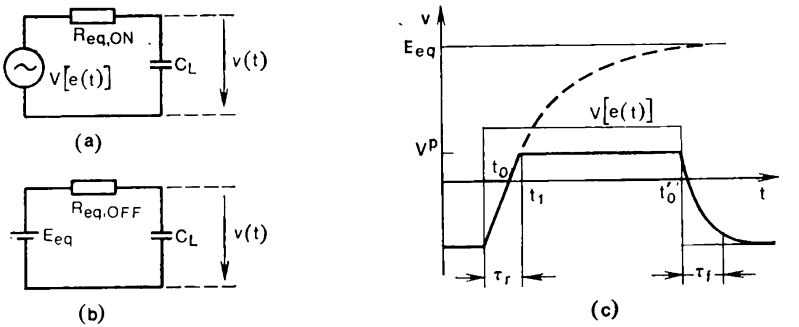


Fig. 4.7

At time t_1, the diode D_2 is rendered conducting. The rise time τ_r can be found from the equation

$$v(t_0 + \tau_r) = V^p$$

whence

$$\tau_r = -C_L R_{eq,\,OFF} \ln\left[1 - \frac{V_m}{E_{eq} - v(0)}\right] \tag{4.9}$$

From a comparison of Eqs. (4.9) and (4.8) it is seen that, in contrast to a parallel limiter, the rise time of a series-parallel limiter is independent of the input signal amplitude. The rise time can be controlled by varying the supply voltage and, as a consequence, the current switched between the two diodes. The possibility of obtaining short rise and fall times is a major advantage of series-parallel limiters.

4.3. TRANSISTOR LIMITERS

The switching elements in transistor limiters are the emitter and collector junctions. As a result, a transistor limiter provides a double-peak (positive and negative) limiting action.

Common-emitter limiter (Fig. 4.8*a*). The transfer characteristic of a common-emitter limiter shown in Fig. 4.8*b* has three straight-line segments each of which represents a particular region (or condition) of operation. The horizontal segment represents positive limiting ($e < E^p$), when the transistor is in the OFF state. Referring to the equivalent circuit of a limiter in which the transistor operates at cutoff (Fig. 4.8*c*), the base and collector voltages are given by

$$\begin{aligned} v_B &= [eR_B + E_B(R_s + R_c) + I_{CBO}(R_s + R_c)R_B]/(R_B + R_c + R_s) \\ v &= V^p = V_0^p = (E_C - I_{CBO}R_C)R_L/(R_L + R_C) \end{aligned} \tag{4.10}$$

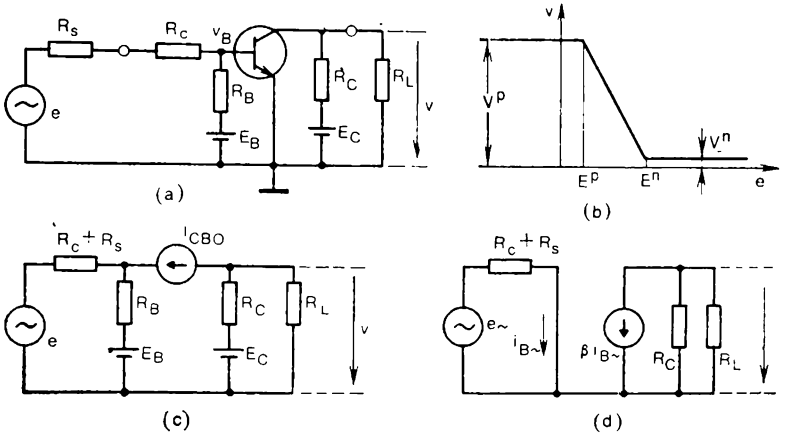


Fig. 4.8

The transistor will remain in the cutoff region so long as $v_B \leqslant e_{EO}$. Solving the equation $v_B = e_{EO}$ for the input voltage gives the positive limiting threshold:

$$E^p = e_{EO} + (R_c + R_s)\,[(e_{EO} - E_B)/R_B - I_{CBO}]$$

The sloping segment of the transfer characteristic represents the amplifying action of the transistor. The gain K can be found from the a.c. equivalent circuit shown in Fig. 4.8*d*:

$$K = v_\sim/e_\sim = -\frac{\beta i_{B\sim}(R_C \parallel R_L)}{e_\sim}$$

$$= -\frac{\beta R_C R_L}{(R_C + R_L)(R_c + R_s)}$$

In negative limiting ($e > E^n$), the transistor is at saturation, and the collector voltage is close to zero, that is,

$$V^n = V_0^n \approx 0$$

On setting in Eq. (4.3) $K^p = K^n = 0$ and $V_0^n \ll V_0^p$, we can express E^n in terms of E^p, V^p, and K as follows:

$$E^n = E^p + V^p/K = e_{EO} + (R_c + R_s)\left(\frac{e_{EO} - E_B}{R_B} - \frac{I_{CBO}}{\alpha} + \frac{E_C}{\beta R_C}\right)$$

To minimize the effect of variations in I_{CBO} on the limiter parameters, it will be well-advised to choose R_B such that

$$(e_{EO} - E_B)/R_B \gg I_{CBO}$$

Then the terms containing I_{CBO} can be removed from the expressions for E^n and E^p.

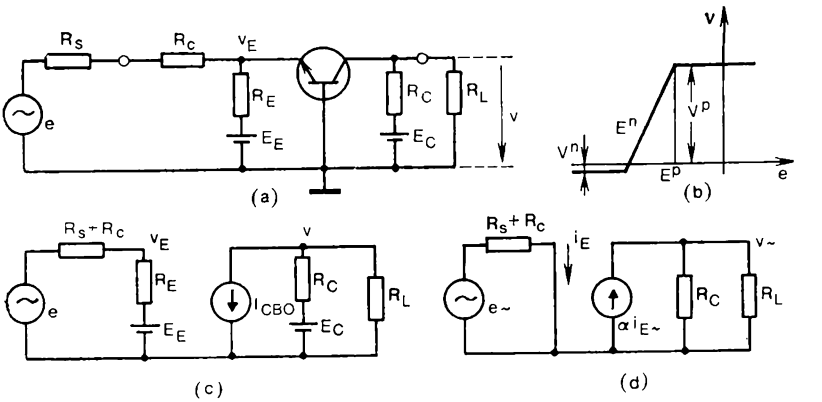


Fig. 4.9

Common-base limiter (Fig. 4.9). During positive limiting ($e > E^p$) transistor is OFF. Using the simplified piecewise-linear transistor model, we derive the equivalent circuit of Fig. 4.9*c* from which it follows that

$$v_E = [eR_E + E_E(R_s + R_c)]/(R_E + R_c + R_s)$$

$$v = V^p = V_0^p = (E_C - I_{CBO}R_C)\,R_L/(R_C + R_L)$$

The transistor will remain at cutoff so long as $v_E > -e_{EO}$. Therefore,

$$E^p = -e_{EO} - (e_{OE} + E_E)(R_c + R_s)/R_E$$

The sloping portion of the transfer characteristic applies when the transistor is amplifying the signal. Using the a.c. equivalent circuit of Fig. 4.9*d*, the gain is found to be

$$K = v_\sim/e_\sim = \frac{\alpha i_{E_\sim}(R_C \parallel R_L)}{e_\sim} = \frac{\alpha R_L R_C}{(R_L + R_C)(R_c + R_s)}$$

During negative limiting ($e < E^n$), the transistor is at saturation, and the collector voltage is controlled by $-e_{EO}$, so that

$$V^n = V_0^n = -e_{CO}$$

On setting $K^n = K^p = 0$ in Eq. (4.3) and recalling that $e_{CO} \approx e_{EO}$, we get

$$E^n = E^p - (V^p - V^n)/K = -[e_{EO} + e_{EO}(R_c + R_s)(1/R_E + 1/R_C + 1/R_L) + \frac{(E_C - I_{CBO}R_C)(R_c + R_L)}{\alpha R_C} + E_E(R_c + R_s)/R_E]$$

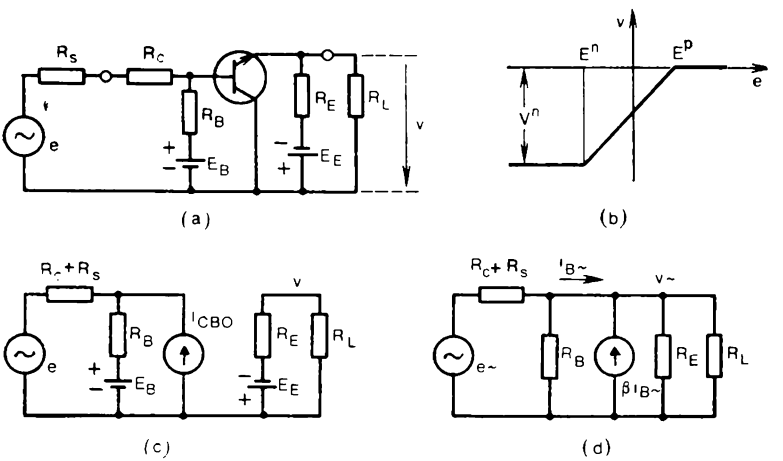


Fig. 4.10

The effect of I_{CBO} on the performance of the limiter can be minimized by choosing R_C such that $R_C I_{CBO} \ll E_C$. The stability of a common-base limiter is higher than that of a common-emitter circuit, because the relative change in α is $1/(1+\beta)$ times the variations in β.

Common-collector limiter (Fig. 4.10). In negative limiting ($e < E^n$), the transistor is nonconducting, and the equivalent circuit of the limiter takes the form shown in Fig. 4.10. Hence, the expression for v_B is the same as Eq. (4.10), and the collector voltage is given by

$$v = V^n = -E_E R_L/(R_L + R_E)$$

For the transistor to be at cutoff, it is required that

$$v_B - V^n < e_{EO}$$

whence

$$E_-^n = -E_E R_L/(R_L + R_E) - (E_E + E_B + I_{CBO} R_B)\frac{R_C + R_s}{R_B} + e_{EO}$$

The sloping portion of the transfer characteristic applies when the transistor is amplifying the input signal. From the a.c. equivalent circuit of Fig. 4.10*d*, the gain is found to be

$$K = v_\sim/e_\sim = \frac{R_B}{R_B + R_C + R_s}\left[1 + \frac{1}{1+\beta}\frac{(R_C + R_s) \parallel R_B}{R_L \parallel R_E}\right]^{-1} \quad (4.11)$$

At $\beta \gg 1$, Eq. (4.11) takes on a simpler form

$$K \approx R_B/(R_B + R_C + R_s)$$

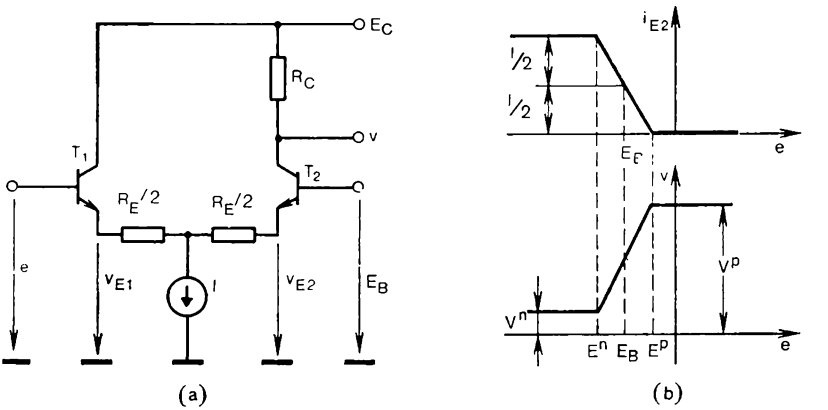


Fig. 4.11

During positive limiting ($e > E^p$), the transistor is at saturation, and the output voltage is practically zero:

$$v = V^p - V_0^p \approx 0$$

Hence,

$$E^p = E^n - V_0^n/K = E_E \frac{R_L (R_C + R_s)}{R_B (R_L + R_E)} - (E_E + E_B - e_{OE} + I_{CBO} R_B) \frac{R_C + R_s}{R_B} + e_{EO}$$

In stability, a common-collector limiter is close to a common-base circuit. It suffers from a disadvantage that $K \leqslant 1$.

Emitter-coupled limiter (Fig. 4.11*a*). In contrast to the previous circuits, this configuration uses two transistors which form a differential amplifier stage. Among the advantages of this circuit are high stability of performance and high speed of response. The high speed is secured because the double-peak (positive-negative) limiting action is obtained by utilizing the switching properties of only the emitter junctions. The collector junctions remain in the OFF state both when the signal is passed and when it is limited (the transistors operate in the unsaturated region). As a corollary, the limiter is free from the storage delay associated with the accumulation and removal of excess minority-carrier charge in and from the base and collector regions.

Consider the relations defining the transfer characteristic of an emitter-coupled limiter. The emitter currents of the transistor are each the sum of two components, one being due to the current generator I, and the other to the voltage difference $v_{E1} - v_{E2}$. In the

circumstances, the following relations apply:

$$i_{E1} = 0.5I + (v_{E1} - v_{E2})/R_E \qquad i_{E2} = 0.5I - (v_{E1} - v_{E2})/R_E$$

Neglecting the resistance of the conducting emitter junction because it is negligible in comparison with R_E, we may take it that in the signal transmission region,

$$v_{E1} - v_{E2} = e - E_B$$

because the base-to-emitter voltages cancel out.

The emitter junction of T_1 is reverse-biased at $i_{E1} = 0$, and that of T_2, at $i_{E2} = 0$. Solving these equations for the input signal gives the threshold voltages

$$E^n = E_B - 0.5IR_E, \quad E^p = E_B + 0.5IR_E$$

When T_1 is OFF, $i_{E2} = I$, and

$$v = V^n = V_0^n = E_C - I_{CBO}R_C - I\alpha R_C \approx E_C - IR_C$$

When T_2 is OFF, $i_{C2} = I_{CBO}$, so

$$v = V^p = V_0^p = E_C - I_{CBO}R_C \approx E_C$$

On setting $K^n = K^p = 0$, we obtain from Eq. (4.3)

$$K = (V_0^p - V_0^n)/(E^p - E^n) \approx R_L/R_E$$

For the correct operation of an emitter-coupled limiter, it is essential that the conducting transistor is in the unsaturated condition. This in turn requires that $E_C - IR_C > E_B$. The static characteristics of an emitter-coupled limiter are shown in Fig. 4.11*b*.

Since the circuit of an emitter-coupled limiter is symmetrical, the input signals can simultaneously be applied to the bases of both transistors (Fig. 4.12*a*), and the output signal v (t) is the difference of two voltages, $e_1 - e_2$. The input e_1 is referred to as the noninverting input, and the input e_2 as the inverting input. By holding e_1 constant, we obtain a transfer characteristic, v (e_2) (Fig. 4.12*b*) which differs from that examined earlier in having a negative gain, K, in the signal transmission region. As e_2 rises, the instant is reached when T_2 is saturated, and from that instant on the output voltage

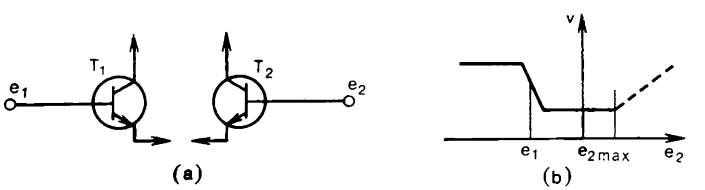


Fig. 4.12

follows any variations in the input (the dashed line in Fig. 4.12*b*). The point where T_2 moves into the saturation region defines the maximum safe level for the input signal, $e_{2,\,max}$.

4.4. SWITCHING MODE OF POSITIVE-NEGATIVE (DOUBLE-PEAK) LIMITERS

The transistor switches discussed in previous section are widely used in logic gates, flip-flops, multivibrators, and other pulse circuits. In most cases, stable states exist only in signal limiting, when the output voltage is equal to the positive or negative limiting level. This is known as the switching mode of a positive-negative limiter.

In the switching mode, a limiter can be treated as a switch with two states differing in output voltage. In the region where the applied signal is allowed to pass through without limiting (it is represented by the sloping portion of the transfer characteristic), the switch is in an indeterminate state. Therefore, the switching behaviour of the limiter is improved as this region is made narrower. The ideal switching performance is shown by circuits whose transfer characteristic is a step response function. All existing types of switches used in pulse circuits have only a quasi-step response approaching the ideal with a varying degree of accuracy.

The current trend in the Soviet Union is to use an integrated-circuit differential opamp as a component part of pulse circuits. The typical transfer characteristic of such amplifiers is shown in Fig. 4.13. Owing to the high gain in the transmission region (about 10^3 to 10^5), a differential d.c. amplifier can legitimately be treated as an ideal switch. The differential input extends the functional capabilities of the circuit. In more detail, the static and dynamic behaviour and circuit design of differential d.c. amplifiers are examined in Sec. 13.5.

Improvement of switch performance by a differential opamp. Consider the switching circuit shown in Fig. 14.14*a*. The shape of its *V-I* characteristic, $i_{sw} = f(v_{sw})$, shown in Fig. 4.14*b*, is governed by the nonlinear behaviour of both the diode and the opamp. When v_{sw} is positive, the diode is conducting, and to a first approximation the output voltage of the opamp, v, is approximately equal to V_f.

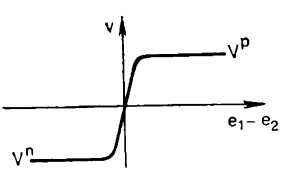


Fig. 4.13

At the same value of i_{sw}, the input voltage of the opamp, v_{sw}, is approximately equal to V_f/K_0. In other words, the residual parameters of the closed switch are $1/K$th of their original value. When v_{sw} is negative, the diode is at cutoff, and the resistance of the switch remains $1/K_0$ part of the diode's reverse resistance until the opamp moves into the limiting region

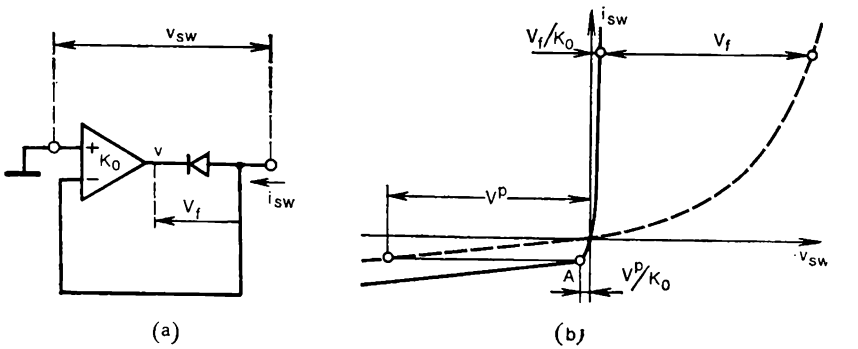


Fig. 4.14

(point A). As $|\, v_{sw} \,|$ keeps rising, the output voltage of the opamp no longer decreases, and the incremental (dynamic) resistance of the switch abruptly rises to become equal to the reverse resistance of the diode. As is seen, the switch built around a differential opamp has the diode type of V-I characteristic, but this characteristic has a more pronounced breakpoint than that of a crystal diode.

A similar circuit can be constructed using a reverse-biased diode or a transistor.

Switches incorporating differential opamps are utilized in precision limiters, clamping circuits (see Sec. 4.5), and some other applications.

4.5. CLAMPING CIRCUITS

Clamping circuits, also known as d.c. restorers, serve to clamp the base or top of a pulse signal to some specified level when voltage or current pulses are passed through circuits containing d.c. blocking capacitors or transformers.

For its operation, a clamping circuit depends on the resistors of an RC-network that are switched to provide a suitable time constant or the charge and discharge circuits of a d.c. blocking capacitor (or the rise and fall circuit of the magnetizing current in a transformer) during and between the pulses being transmitted.

Consider how voltage pulses are clamped at the output of d.c. blocking RC-networks (Fig. 4.15*a*) where the switch is closed between pulses and open during the existence of a rectangular pulse (Fig.4.15*b*).

The time constant for the OFF and ON states of the switch is respectively given by $\tau_0 = CR$ and $\tau_1 = C\,(R \parallel r_{sw}) \approx Cr_{sw}$ (for $R \gg r_{sw}$).

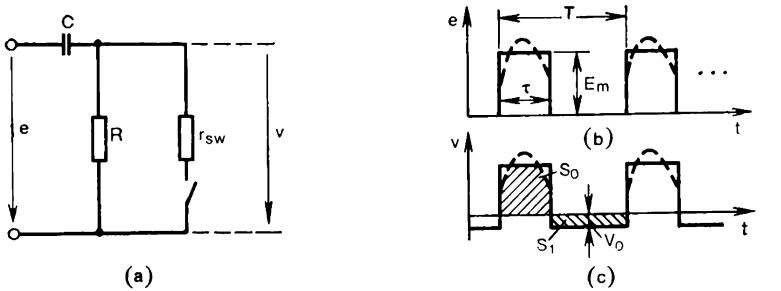


Fig. 4.15

The condition for charge balance, as defined by Eq. (2.3a) for a state of dynamic equilibrium may be interpreted as implying that the ratio of the "positive" to the "negative" area of the output signal is such that

$$A_1/A_0 = R/(R \parallel r_{sw}) = \tau_0/\tau_1 \tag{4.12}$$

Suppose that in order to keep to a minimum the distortion in the signal waveform the value of C is chosen so large that

$$\tau_1 \gg T - \tau \text{ and } \tau_0 \gg \tau \tag{4.13}$$

The applicable signal waveform is shown in Fig. 4.15*c*. Approximating the shaded areas A_0 and A_1 by rectangles and using Eq. (4.12), we find the dynamic bias to be

$$V_0 = E_m\,[1 + (\tau_0/\tau_1)\ (T/\tau - 1)]^{-1} \tag{4.14}$$

The above relation holds true also when the pulse waveform is other than rectangular (the dashed line in Fig. 4.15*b*). The value of E_m is then defined as the amplitude of a rectangular pulse equal in area and duration to the input pulse.

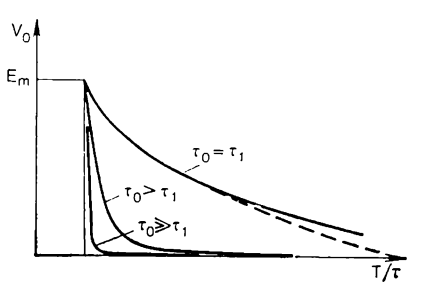


Fig. 4.16

Figure 4.16 shows plots of the dynamic bias V_0 as a function of the period-to-duration ratio, T/τ, for a pulse train. These plots are described by Eq. (4.14), and constructed for several relative values of τ_0 and τ_1. The case of $\tau_0 = \tau_1$ applies to a linear (nonswitched) circuit. There is a considerable range of pulse period-to-duration ratios where the value of V_0 is strongly affected.

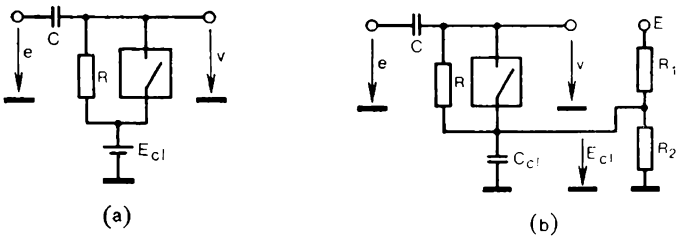


Fig. 4.17

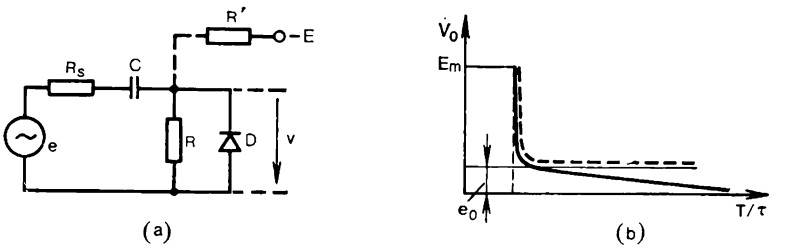


Fig. 4.18

As the ratio increases with increasing T, an instant is reached when the inequality $\tau_1 \gg T$ is no longer satisfied, and V_0 falls at a faster rate (the dashed line in Fig. 4.16) than it should in accordance with Eq. (4.14).

The clamping action takes place at $\tau_0 \gg \tau_1$. Even a slight deviation of T/τ from unity will suffice to make V_0 practically equal to zero. At low values of τ_1, the inequality $\tau_1 \gg T$ is no longer satisfied, and V_0 decreases at a markedly faster rate than it should be expected from Eq. (4.14). In this way, the clamping circuit maintains V_0 at a constant value (zero value in our case) over the entire range of likely values for the pulse period-to-duration ratio.

Clamping to a level other than zero is provided by the circuit shown in Fig. 4.17*a* where a source of clamping voltage, E_{cl}, is placed in series with the resistor of the d.c. blocking network. If the efficiency of the clamping action is not to be impaired, the internal source resistance must be small in comparison with r_{sw}. In practice, the desired clamping voltage is derived from a voltage divider shunted by a high-value capacitor ($C_{cl} \gg C$, Fig. 4.17*b*) rather than from a separate source. The divider presents a low impedance to the a.c. components. In the circuit in question, this is a sufficient measure, because the direct component of current through the source of E_{cl} is zero.

Practical clamping circuits differ in the type of switching element used and in the manner it is operated.

Figure 4.18*a* shows a diode clamping circuit which uses negative clamp pulses. The diode turns on when the output signal becomes negative and turns off when the output signal is positive (see Fig. 4.15*b*). Including the internal resistance of the signal source, the time constants of the d.c. blocking circuit are $\tau_0 = C\,(R_s + R \parallel r_r)$ and $\tau_1 = C\,(R_s + R \parallel r_f)$. It is seen that for the clamping action to take place, the values of R and the diode's reverse resistance must be increased, whereas the values of R_s and the diode's forward resistance must be decreased.

Since the forward resistance of the diode begins to increase abruptly at a voltage less than e_0, the real dependence of dynamic bias on the pulse period-to-duration ratio is other than ideal (see Fig. 4.18*b*). Gradual variations in V_0 with the period-to-duration ratio inside the interval from zero to e_0 is often undesirable. The manner in which V_0 varies with T/τ on introducing R' and $-E$ in the circuit is illustrated in Fig. 4.18*b* by the dashed line.

Figure 4.19*a* shows a diode clamping circuit which uses positive clamp pulses. The diode is conducting when the output signal is positive. The response of the circuit to a train of rectangular pulses is shown in Fig. 4.19*b*. The dynamic bias V_0 is defined by the same qualitative relations as before, if we take a negative-going change of signal voltage as the pulse.

There are also clamping circuits which establish the extremities of the signal both above and below some constant reference level. This form of clamping is utilized in colour television, the sweep circuits of oscilloscopes, etc. The state of the switch in such a clamping circuit is controlled by a suitable signal synchronized with the voltage being transmitted. Accordingly, such clamping circuits are known as controlled.

A controlled clamping circuit using a transistor switch is shown along with its time waveform in Fig. 4.20. When no control signal is applied, the transistor is at saturation because its base is returned via R_B to a source of $+E$. Because in a steady state the collector current is zero, saturation is achieved with any base current, $i_B \approx E_B/R_B$. It is desirable, however, to keep the base current high so as to minimize the resistance of the switch in the closed condition.

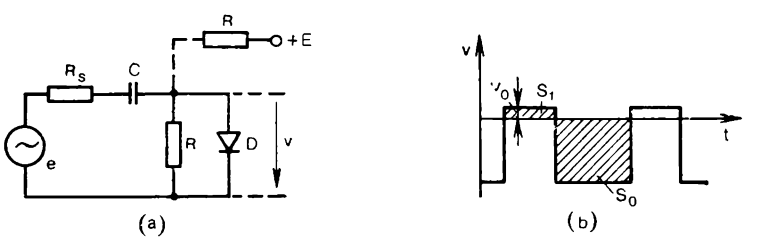


Fig. 4.19

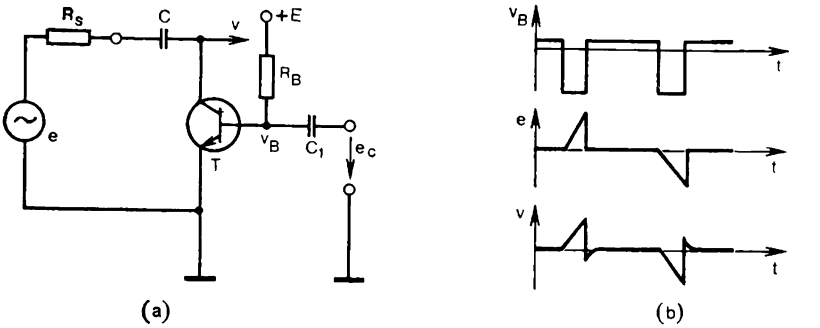


Fig. 4.20

A negative pulse applied to the control input derives the transistor to cutoff. As a result, a pulse of any arbitrary waveshape and polarity can be passed through the d.c. blocking circuit during the existence of the input pulse. Because a decrease in the collector voltage might render the collector junction conducting, it is essential for the correct operation of the circuit that the maximum amplitude of the negative pulses should not exceed that of the control pulses.

Clamping in amplifiers. It is often necessary to use a d.c. blocking *RC*-network at the input of an inverting amplifier. In such cases, it will be well advised to place a clamping switch in the feedback path (Fig. 4.21). Then, according to Eq. (2.32), the time constants for the open and closed switch will be

$$\tau_0 = C\,[R_s + R_{in} \parallel (r_o/K_0)]$$

and

$$\tau_1 = C\,(R_s + R_{in} \parallel (r_c/K_0))$$

For the clamping action to take place ($\tau_0 \gg \tau_1$), it is essential that

$$R_s \ll r_{in},\ r_o/K_i r_c \ll r_o$$

where r_{in} is the input resistance of the amplifier, r_o is the resistance of the switch when open, and r_c is the resistance of the switch when closed.

In the above circuit, the signal is clamped to E_c. In practice, however, the clamping is not perfect; the output voltage rises slightly above E_c and the waveshape is a somewhat distorted version of the input. These discrepancies depend on the zero drift of the amplifier and the residual voltage across the switch.

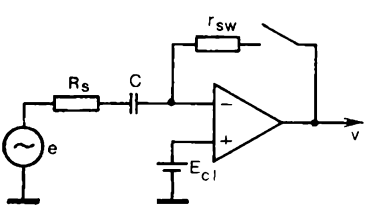


Fig. 4.21

REVIEW QUESTIONS AND PROBLEMS

4.1. The parameters of a positive limiter are: $K = 1$, $V_0 = 0$, $K^{\text{p}} = 0.1$, and $V_0^{\text{p}} = 5$ V. Determine E^{p} and V^{p} and plot the waveform of the output signal, assuming that the input signal is a sine wave with an amplitude of 10 V.

4.2. The transfer characteristic of a limiter is specified by the function $v = \max(e, 0.1e)$. Identify the type of limiter, plot its transfer characteristic and the waveform of the output signal for a sine-wave input signal with an amplitude of 5 V.

4.3. Solve Problem 4.2 for $v = \min[\max(e, 0.1e), 2 + 0.1e]$.

4.4. List the types of diode limiters you know and point out how they differ.

4.5. Using the piecewise-linear diode model, determine the transfer characteristic for a double-peak (positive-negative) parallel diode limiter. Derive expressions for K, K^{p}, K^{n}, E^{p} and E^{n}, with allowance for the internal resistance of the signal source and the load resistance.

4.6. Solve Problem 4.5 for a series and a series-parallel diode limiter.

4.7. Develop the circuit of a double-peak (positive-negative) series diode limiter with $E^{\text{n}} = 10$ V and $E^{\text{p}} = 30$ V, driven by a source of $E = 100$ V. Carry out the calculations, using the ideal diode model.

4.8. In the parallel diode limiter of Fig. 4.2*a*, the diode is a silicon Zener diode with a breakdown voltage of 5.6 V. Plot an approximate transfer characteristic of the limiter for $E = 2$ V.

4.9. Plot the transfer characteristic for the opamp limiter of Fig. 4.2*k*, with the bias sources omitted ($E_1 = E_2 = 0$), and the diodes being the Soviet-made type Д219 silicon diodes (see Tables 3.1 and 3.2).

4.10. Plot the transfer characteristic for the opamp limiter of Fig. 4.2*j* in which $E = 0$ and the diode is a silicon Zener diode with a breakdown voltage of 4.7 V.

4.11. Using a series-parallel limiter, develop diode switches deriving from two input voltages e_1 and e_2 an output voltage equal to the maximum and the minimum one. Construct similar circuits with a larger number of inputs.

4.12. The parallel diode limiter of Fig. 4.2 with $R = 1$ kΩ, and $E = 5$ V is driven by an input rectangular pulse with a duration of 10 μs and an amplitude of 12 V, assuming that in the no-signal state $e = -2$ V. Determine the waveform of output voltage, if the limiter is loaded into $C_L = 1$ nF, and the internal resistance of the signal source is 50Ω. Using the ideal diode model, calculate the rise and fall times of the output pulse.

4.13. Solve Problem 4.12 for the series limiter of Fig. 4.2*b*.

4.14. Solve Problem 4.12 for the series-parallel limiter of Fig. 4.2*h* in which $E = 20$ V and $R = 1$ kΩ.

4.15. Solve Problem 4.12 for the double-peak (positive-negative) series-parallel limiter of Fig. 4.2*i* in which $E_1 = E_2 = 20$ V, $R_1 = 1$ kΩ, and $R_2 = 5$ kΩ.

4.16. Develop the circuit of a common-emitter transistor limiter delivering an output signal with an amplitude of 5 V and operating at limiting thresholds $E^{\mathrm{p}} = 0.5$ V and $E^{\mathrm{n}} = 1.5$ V. The transistor parameters are: $\beta = 40$, $e_{\mathrm{EO}} = 0.6$ V, $I_{\mathrm{CBO}} = 0$, and $I_{\mathrm{C,\,max}} = 10$ mA.

4.17. Plot the *V-I* characteristic for the switch of Fig. 4.18*a* and determine its parameters, using the piecewise-linear diode model in which $r_{\mathrm{r}} = 10^6\,\Omega$, $r_{\mathrm{f}} = 10\,\Omega$, and $e_0 = 0.6$ V. The amplifier gain is $K_0 = 10^4$, and its output signal is limited at ± 10 V.

4.18. The diode clamping circuit using negative clamp pulses (Fig. 4.21) accepts a periodic sequence of positive pulses with an amplitude E and a pulse period-to-duration ratio Q. Using the piecewise-linear diode model corresponding to the equivalent circuit of Fig. 3.6*b* and assuming that $T \ll C\,(R_s + r_{\mathrm{f}})$, express the charging and discharge currents of C in terms of the output voltage; formulate the condition for the diode to conduct between pulses; evaluate the relation $V_0\,(Q)$ on the basis of the condition for the balance of capacitor charging and discharge currents: $i_{\mathrm{ch}} = i_{\mathrm{dis}}\,(Q - 1)$.

5 LOGIC GATES

5.1. SWITCHING FUNCTIONS

The rationale of digital data processing is that data can be presented as an ordered collection of binary variables $x_1, x_2, \ldots, x_n$, each of which can take on any one of two values, a 0 or a 1, often called a *logic 0* and a *logic 1*. A 1 can be treated as the magnitude of a number in binary notation or as an indication that an event has occurred within or at an input to a given circuit or device. This event may be the closure of a switch, reversal in the polarity of a time-varying voltage, a throw-over of a toggle on a control panel, closure of a limit switch in an electromechanical system, arrival of an echo signal in a radar, etc.

A basic operation in digital data processing is to implement functions $y = f(x_1, x_2, \ldots, x_n)$ which assign a value of 0 or 1 to each of the 2^n possible combinations of the variables $x_1, x_2, \ldots, x_n$. The rules according to which this is done are known as *switching functions.*

As an example, consider the arithmetic operation of addition of quantities A and B expressed as binary numbers: $a_1, a_2, \ldots, a_N$ and $b_1, b_2, \ldots, b_N$. Finding the sum $C = A + B$ reduces to finding a binary number $c_1, c_2, \ldots, c_N$ each digit of which is the switching function of the binary variables $a_1, a_2, \ldots, a_N$ and $b_1, b_2, \ldots, b_N$. That is,

$$c_i = f_i(a_1, a_2, \ldots, a_N; b_1, b_2, \ldots, b_N)$$

where $i = 1, 2, \ldots, N'$.

One way to specify a switching function is by what is known as a *truth table.* The truth table for a switching function exhausts the combinations of values of the arguments in the leftmost column, and the values of the function corresponding to each of the possible 2^n combinations of values of the variables in the remaining columns. A single truth table may list several functions of the same arguments. As an example, consider the switching functions specified by a truth table, Table 5.1, and implementing the addition of two-digit numbers A and B. The sum is a three-digit number C. The truth table specified the switching functions $c_1(a_1, a_2; b_1, b_2)$, $c_2(a_1, a_2; b_1, b_2)$, and $c_3(a_1, a_2; b_1, b_2)$. The validity of the truth table can be verified by carrying out the operation of addition. As an example, for line 11

Table 5.1. Truth Table for the Operation of Addition

Line No.	a_1	a_2	b_1	b_2	c_1	c_2	c_3	Line No.	a_1	a_2	b_1	b_2	c_1	c_2	c_3
0	0	0	0	0	0	0	0	8	1	0	0	0	0	1	0
1	0	0	0	1	0	0	1	9	1	0	0	1	0	1	1
2	0	0	1	0	0	1	0	10	1	0	1	0	1	0	0
3	0	0	1	1	0	1	1	11	1	0	1	1	1	0	1
4	0	1	0	0	0	0	1	12	1	1	0	0	0	1	1
5	0	1	0	1	0	1	0	13	1	1	0	1	1	0	0
6	0	1	1	0	0	1	1	14	1	1	1	0	1	0	1
7	0	1	1	1	1	0	0	15	1	1	1	1	1	1	0

we have

$$\begin{array}{r} 10 \\ +\,11 \\ \hline 101 \end{array} \equiv \begin{array}{r} 2 \\ +\,3 \\ \hline 5 \end{array}$$

Another way to specify a switching function is by what is known as a *switching expression*, describing the order of precedence of elementary operations to be performed on the arguments of the function. The three basic operations (the *basic set*) of switching algebra are the AND operation (logic multiplication or conjunction), the OR operation (logic addition or disjunction), and the NOT operation (complementation or negation).

The logic product (conjunction) of the variables $x_1, x_2, \ldots, x_N$ is a binary variable y which takes on a value 1 if and only if all operands (multiplicands) take on a value 1. In other words, for the product to be zero, it will suffice for one of the operands to be zero. The logic product is denoted $y = x_1x_2 \ldots x_N$, or $y = x_1 \wedge x_2 \wedge \ldots \wedge x_N$ (read as "x_1 AND x_2 AND ... AND x_N").

The logic sum (disjunction) of the variables $x_1, x_2, \ldots, x_N$ is a binary variable y which takes on a value zero if and only if all the operands are zero. In other words, for a logic sum to be unity, it will suffice for at least one of the operands to be unity. The logic sum is denoted $y = x_1 + x_2 + \ldots + x_N$ or $y = x_1 \vee x_2 \vee \ldots \vee x_N$ (read as "x_1 OR x_2 OR ... OR x_N").

The logic complement of the variable x is a binary variable y which takes on a value 1 (zero) if and only if x is zero (unity). Logic negation is denoted $y = \bar{x}$ (read as "NOT x").

The three operations listed above form the basis of switching algebra which is a special case of a more general class of algebra known as *Boolean algebra.*

As already noted, a switching function can be specified by a switching expression, and this can be done in more than one way, by equivalently transforming one expression to another. This is

done according to a few basic rules or axioms. Some of them are defined below.

(1) Involution (double negation):

$$\bar{\bar{x}} = x$$

(2) Commutation:

$$xy = yx$$

$$x + y = y + x$$

(3) Associativity:

$$x(yz) = (xy)z$$

$$x + (y + z) = (x + y) + z$$

(4) Distribution:

$$x(y + z) = xy + xz$$

$$x + yz = (x + y)(x + z)$$

(5) Multiplication:

$$xx = x$$

$$x0 = 0$$

$$x1 = x$$

$$x\bar{x} = 0$$

(6) Addition:

$$x + x = x$$

$$x + 0 = x$$

$$x + 1 = 1$$

$$x + \bar{x} = 1$$

(7) De Morgan's theorem:

$$\overline{x + y} = \bar{x}\bar{y}$$

$$\overline{xy} = \bar{x} \times \bar{y}$$

As an example, consider the following switching expressions:

$$c_3 = a_2\bar{b}_2 + \bar{a}_2 b_2$$

$$c_2 = (a_1\bar{b}_1 + \bar{a}_1 b_1)\,\overline{a_2 b_2} + \overline{(a_1\bar{b}_1 + \bar{a}_1 b_1)\,a_2 b_2}$$

$$c_1 = a_1 b_1 + a_2 b_2 (a_1 + b_1)$$

They define the operation of addition on two-digit binary numbers. On substituting the values of the arguments, it is an easy matter to check that the values of the functions specified by switching expressions are the same as specified by a truth table.

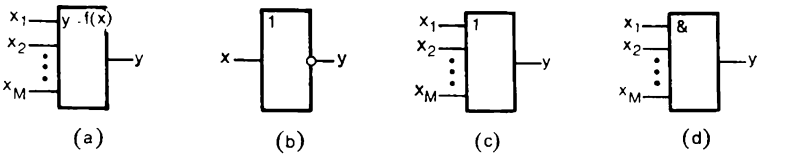


Fig. 5.1

Switching functions are implemented by electronic circuits known as logic gates. As a rule, a logic gate has $M \geqslant 1$ inputs and one output. In Soviet practice, a logic gate is diagrammatically shown as a rectangle enclosing the switching function to be implemented (Fig. 5.1*a*). The OR operation is denoted by a 1, the AND operation by an ampersand (&), and the NOT operation by a circle at the output (input) of the rectangle. The circuit performing the operation of negation (complementation) is called the NOT gate or inverter (Fig. 5.1*b*), the circuit performing the operation of logic addition is called the OR gate (Fig. 5.1*c*) and the circuit performing the operation of logic multiplication is called the AND gate (Fig. 5.1*d*).

In the general case, a switching function is mechanized by connecting together several logic gates each of which performs one of the basic functions, namely the AND operation, the OR operation, and the NOT operation which between them form a complete set of logic gates.

The connection of the basic gates into a switching circuit is uniquely related to the switching function it mechanizes. As an example, consider the switching function

$$c_1 = a_1b_1 + a_2b_2\,(a_1 + b_1)$$

Let us re-write it as a set of equations

$$x_1 = a_1b_1$$
$$x_2 = a_1 + b_1$$
$$x_3 = x_2a_2b_2$$
$$c_1 = x_1 + x_3$$

each of which is mechanized by a particular gate, and the switching function as a whole is implemented by the circuit shown in Fig. 5.2.

The number of logic-gate types used in logical design can be minimized by using the logical gates which are in themselves complete, although each performs only one logical operation. Most convenient of them are the following sets:

—the NOT-OR (NOR) gate implementing the Pierce function:

$$y = x_1 \downarrow x_2 \downarrow \ldots \downarrow x_m = \overline{x_1 + x_2 + \ldots + x_m}$$

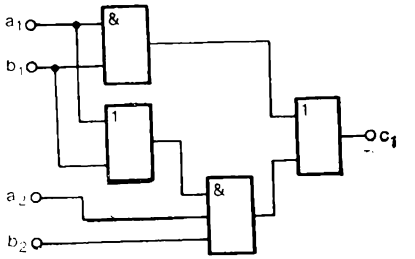


Fig. 5.2

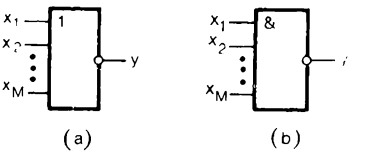


Fig. 5.3

where $m = 1, 2, \ldots, M$. The graphic symbol for a NOR gate is shown in Fig. 5.3*a*;

—the NOT-AND (NAND) gate implementing the Sheffer function:

$$y = x_1/x_2/\ldots/x_m = \overline{x_1 + x_2 + \ldots + x_m}$$

where $m = 1, 2, \ldots, M$. The graphic symbol for a NAND gate is shown in Fig. 5.3*b*.

A complete set formed by an AND, an OR and a NOT gate can be simplified to a NOR or a NAND set, using axioms 1 and 7:

$$x_1 x_2 \ldots x_m = \overline{\overline{x_1 x_2 \ldots x_m}} = \overline{x_1/x_2/\ldots/x_m} = \bar{x}_1 \downarrow \bar{x}_2 \downarrow \ldots \downarrow \bar{x}_m$$

$$x_1 + x_2 + \ldots + x_m = \overline{x_1 + x_2 + \ldots + x_m}$$

$$= \bar{x}_1/\bar{x}_2/\ldots/\bar{x}_m = \overline{x_1 \downarrow x_2 \downarrow \ldots \downarrow x_m}$$

Theoretical studies into digital circuits often resort to an algebraic basic set involving modulo 2 multiplication and modulo 2 addition. Modulo 2 multiplication is analogous to logical multiplication (conjunction). Modulo 2 addition is carried out to the following rule: the modulo 2 sum of binary variables $x_1, x_2, \ldots, x_M$ (denoted as $x_1 \oplus x_2 \oplus \ldots \oplus x_M$) takes on a value of 1, if the number of operands having a value of 1 is odd. For a fixed number M, this rule can be stated in terms of switching algebra. Taking two terms, we may write thus:

$$x_1 \oplus x_2 = x_1\bar{x}_2 + \bar{x}_1 x_2$$

5.2. BASIC CHARACTERISTICS OF LOGIC GATES

Physical presentation of binary variables. As a rule, a parameter equivalent to the value of a binary variable is the output voltage of a logic gate. If this parameter is a d.c. voltage, we have a d.c. or *level logic gate*. For the values of logic variables to pass down a digital network, the stages in the latter must be conductively (resistively) coupled. In the general case, the coupling networks may use

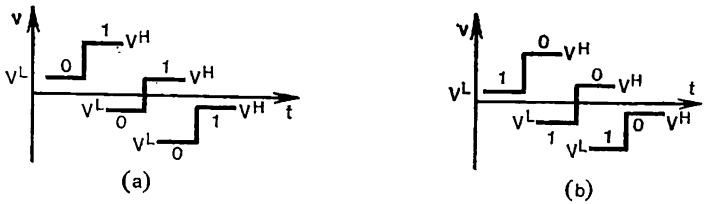


Fig. 5.4

reactances (inductors and capacitors), but they will play an auxiliary role. Level logic gates without reactances are adapted to microelectronic technology best of all.

With level logic gates, the two values of a binary variable are represented by two voltage levels, low (V^L) and high (V^H) (Fig. 5.4)*. In positive logic, a value of unity is represented by a high level, V^H; in negative logic, a value of unity is represented by a low level, V^L.

In another type, known as *pulse logic gates*, the value of the output logic variable is defined over the duration of a clock pulse applied to a separate (clock) input of the gate. A logic 1 is represented by a pulse at the gate output (Fig. 5.5*a*), and a logic zero, by no pulse (Fig. 5.5*b*). Pulse logic gates are now being used on a limited scale.

Consider the performance of d.c. or level logic gates, namely fan-in, fan-out, power dissipation, and switching times.

Fan-in and fan-out. Logic gates are usually connected in various configurations, so that the output of each gate must be able to drive the inputs of many other gates. The number of similar gates that any one gate can drive is known as the *fan-out* of the gate, N. The number of inputs that can be connected to a gate is called the *fan-in* of the gate, M. For most gates, it is equal to the number of physical inputs.

Fan-in and fan-out determine the logic capabilities of a gate. As a rule, an increase in M and N is accompanied by an impairment in reliability, speed, etc. Typical values of M and N range between 4 and 10.

Power dissipation. The power dissipation of a logic gate is a vital parameter because it decides the maximum packing density of gates. The value of power dissipation varies with the state of a gate. Let the power dissipated by a gate in a static state corresponding to a logic 0 be denoted by P^0, the power dissipated in a static state corresponding to a logic 1 by P^1, the power dissipated when the gate goes from 0 to 1 by P^{01}, and the power dissipated as the gate goes from 1 to 0 by P^{10}.

* Some authors denote the two levels as LO and HI, or as L and H, or in some other way.—Translator's note.

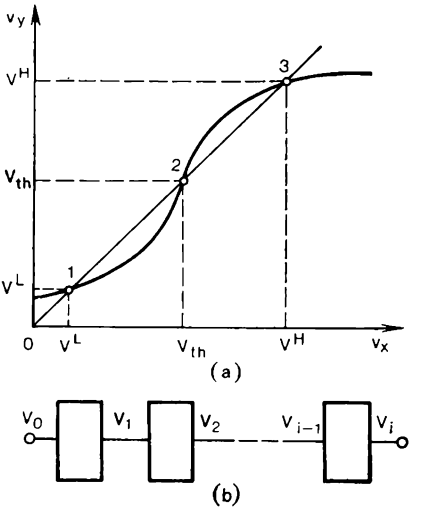


Fig. 5.5

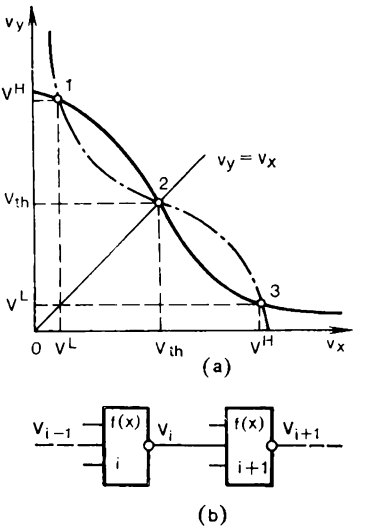


Fig. 5.6

As regards power dissipation, there are two general classes of logic gates. Those in one class dissipate about the same power in a static state and during a change of state. Those in the other draw little power in static states and a large amount of power when changing state. The average power dissipation for the former is $\bar{P} = 0.5\,(P^0 + P^1)$. Apart from the average power, logic gates in the latter class are characterized in terms of the power dissipated by a gate at the maximum switching rate.

Static transfer characteristic of a logic gate. The term 'static transfer characteristic' of a logic gate refers to the dependence of its output voltage v_y on the slowly varying input voltage v_x. There may be a noninverting transfer characteristic (Fig. 5.6*a*) and an inverting (complementing) transfer characteristic (Fig. 5.6*b*). The former is mechanized by noninverting gates (such as AND and OR gates), and the latter by inverting (complementing) gates (NOT, NAND, and NOR gates). With a noninverting transfer characteristic, the changes in v_x and v_y have the same sign, whereas with an inverting transfer characteristic the signs are opposite.

A fundamental property of logic gates, stemming from their transfer characteristics and setting them apart from other switching circuits, is the reproduction of logic levels. More specifically, this property means that if an input of a logic gate is a combination of voltage recognized by the gate as a combination of binary variables, the output of the gate must be a voltage recognized by the next gate

analogous to the preceding one as a logic 0 or a logic 1. This property must be retained under conditions of noise and spread in parameters from sample to sample, with temperature and time.

For an insight into the relation between the shape of the transfer characteristic and the reproduction of logic levels, let us trace the propagation of a signal through a chain of logic gates.

The propagation of a signal through a chain of noninverting gates shown in Fig. 5.5*b* is represented by a sequence of voltages $V_1, V_2, \ldots, V_i$, for a given input voltage V_0. Suppose that the chain is formed by identical gates with a transfer characteristic crossing the line of unity gain, $v_y = v_x$, at three points: *1*, *2* and *3* (see Fig. 5.5*a*). The coordinates of the points are V^{L}, V_{th}, and V^{H}. Because the points of intersection lie on the $v_y = v_x$ line, the voltages V^{L}, V_{th}, and V^{H} applied to the input of the chain will pass through it unaltered. Since the gain of the gates at points *1* and *3* is less than unity, small accidental deviations of the transmitted voltage from V^{L} or V^{H} will be damped out. In contrast, the gate gain at point *2* is greater than unity, so the minute deviations of V_0 from V_{th} will be boosted as the signal advances down the circuit. Depending on the sign of the initial deviation, V_i will tend towards V^{L} or towards V^{H}. Actually, V_{th} is the boundary (threshold) between the regions of convergence towards V^{L} and V^{H}.

Thus, if the transfer characteristic intersects the line of unity gain at three points, a gate will reliably reproduce two voltage levels, V^{L} and V^{H}, which may be treated as the low and high logical levels of the transmitted signal. Should the input voltage deviate from V^{L} and V^{H}, the circuit will recognize $v_x < V_{\mathrm{th}}$ as the low logic level, and $v_x > V_{\mathrm{th}}$ as the high logic level.

Let us compare the above transfer characteristic with that which intersects the $v_y = v_x$ line at only one point (Fig. 5.7). On the right of the intersection, $v_y < v_x$, and on the left of the intersection $v_y > v_x$. Therefore, the sequence of voltages $V_0, V_1, \ldots$ will diminish so long as $V_0 > V_{ss}$ (see Fig. 5.7) and increase so long as $V_0 < V_{ss}$. In either case, the sequence $V_0, V_1, \ldots$ tends towards V_{ss}. This implies that, as the signal advances down the circuit, the two voltage levels required to represent a binary variable are mixed. The transfer characteristic in Fig. 5.7 has a portion where the gate gain exceeds unity. This is, however, not enough for the accurate reproduction of the two logic voltage levels. Also the accurate reproduction of logic levels requires that the gain of a noninverting gate be

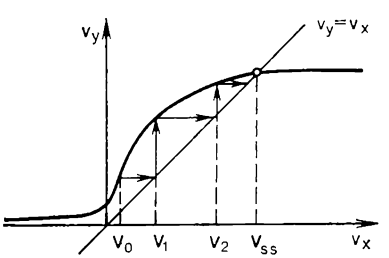


Fig. 5.7

greater than unity at the point where the transfer characteristic intersects the line of unity gain. This requirement arises because the range of values for v_y is limited and there will always be at least two more points of intersection where the gate gain will be less than unity.

Transfer of a signal through a chain of inverting logic gates. The behaviour of a chain of inverting gates stems from that of a chain formed by noninverting gates, if we treat two inverting gates connected in cascade as an elementary noninverting element (see Fig. 5.6*b*). Let us see what transfer characteristic an inverting gate should possess in order that the transfer characteristic of the noninverting element intersect the line of unity gain at three points.

The only point where the transfer characteristic of an inverting gate intersects the $v_y = v_x$ line (point *2* in Fig. 5.6*a*) defines the voltage $v_x = V_{th}$ that can pass through the element unaltered. It follows then that the transfer characteristic of the noninverting element likewise intersects the $v_y = v_x$ line at point *2*, and the gain of that element at that point is equal to the product of the gains of the constituent gates. This is why for two more points of intersection to exist, it will suffice to require that the gain of an inverting gate at point *2* should exceed unity in absolute value.

Because, in absolute value, the gain of a properly operating gate at point *2* must exceed unity, its transfer characteristic intersects its mirror image (the dashed line in Fig. 5.6*a*) at two more points (*1* and *3*) lying symmetrically about the $v_y = v_x$ line. The coordinates of points *1* and *3* define the levels V^L and V^H that will pass through the composite network unaltered.

Thus, if the transfer characteristic intersects its mirror image at three points, a chain of inverting gates will ensure a faithful reproduction of the low (V^L) and high (V^H) logic levels of the input signal. This implies that, should the input voltage v_x deviate from its steady-state levels, a chain of inverting gates will recognize $v_x < V_{th}$ as the low level, and $v_x > V_{th}$ as the high level.

Spread in transfer characteristics. In practice, it happens sometimes that the transfer characteristic of a logic gate is spread over a certain band of values (the shaded area in Figs. 5.8 and 5.9). This occurs because of the spread in component parameters, variations in load, variations in the number of inputs used, etc.

The band of likely transfer characteristics can be defined by specifying the upper and lower bounds. Let the upper bound on the transfer characteristic be denoted by Max F, and the lower bound, by Min F.

The consequences of the spread in the transfer characteristic of a logic gate are as follows:

—logic levels 0 and 1 correspond to nonintersecting intervals of voltages (V^H_{min}, V^H_{max}) and (V^L_{min}, V^L_{max}). The logic levels are correctly reproduced only within these intervals;

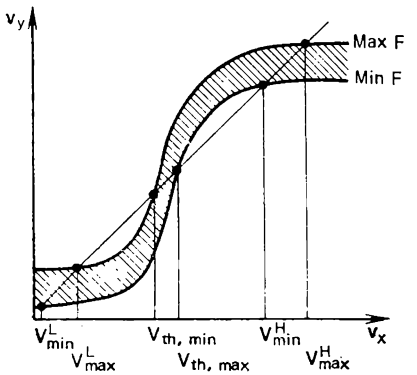


Fig. 5.8

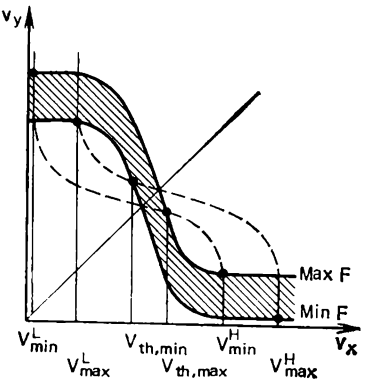


Fig. 5.9

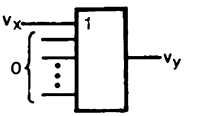
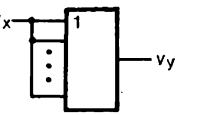
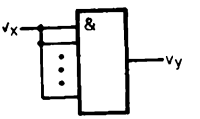
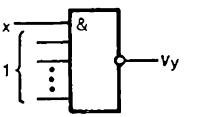
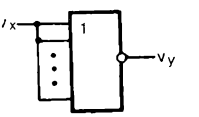
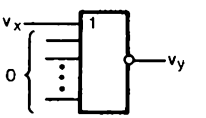
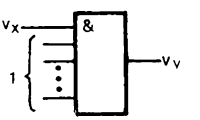
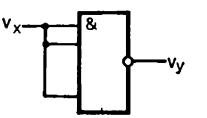
—the threshold voltage of the gate lies within the interval ($V_{\mathbf{th,\,min}}$, $V_{\mathbf{th,\,max}}$). A chain of logic gates recognizes $v_x < V_{\mathbf{th,\,min}}$ as the low voltage level and $v_x > V_{\mathbf{th,max}}$ as the high voltage level. The range of input voltages within the interval ($V_{\mathbf{th,\,min}}$, $V_{\mathbf{th,\,max}}$) defines an indeterminate region where the occurrence of a 0 or a 1 at the chain output is decided by random factors.

For a noninverting logic gate, $V^{\mathbf{H}}_{\mathbf{max}}$, $V_{\mathbf{th,\,min}}$, and $V^{\mathbf{L}}_{\mathbf{max}}$ depend on where the line of unity gain intersects Max F, whereas $V^{\mathbf{H}}_{\mathbf{min}}$, $V_{\mathbf{th,\,max}}$, and $V^{\mathbf{L}}_{\mathbf{th,\,min}}$ are defined by the points where the line of unity gain intersects Min F (see Fig. 5.8). For an inverting logic gate, $V^{\mathbf{H}}_{\mathbf{max}}$, $V_{\mathbf{th,\,min}}$, and $V^{\mathbf{L}}_{\mathbf{max}}$ are defined by the points where Min F intersects a mirror image of Max F, whereas $V^{\mathbf{H}}_{\mathbf{min}}$, $V_{\mathbf{th,\,max}}$, and $V^{\mathbf{L}}_{\mathbf{min}}$ by the points where Max F intersects a mirror image of Min F (see Fig. 5.9).

Should at least one of the above points of intersection be missing on the transfer characteristic of a logic gate, the latter will no longer be capable of faithfully reproducing the logic levels. The decision as to whether a given logic gate can operate properly can only be made after Max F and Min F have been elucidated by calculation or experiment. Therefore, it is important to know how v_x should be applied to a multi-input logic gate in order to obtain the desired Max F and Min F. For positive-logic AND, OR, NAND and NOR gates, the applicable input arrangements are summarized in Table 5.2. They are constructed along the following lines:

—to generate Max F, the signals at the inputs of a noninverting gate must be maintained at a maximum level and at the inputs of an inverting gate at a minimum level;

Table 5.2. Input-Output Arrangements of Logic Gates

Gate type	Min F	Max F
OR		
AND		
NOR		
NAND		

—to generate Min F, a noninverting gate must meet the same requirements as an inverting gate for generating Max F, and vice versa.

When a high level (logic 1) is applied to an OR or a NOR gate to at least one input, the gate goes to a static state. Therefore, the maximum input voltage level is v_x and the minimum voltage level is the logic 0 level at all inputs except one.

AND and NAND gates will go to a static state, if at least one of the inputs accepts a logic 0 level. Therefore, the minimum input voltage is v_x, and the maximum input voltage is a logic 1 level at all inputs except one.

Static noise immunity. The requirements for proper logic level reproduction formulated above define what is necessary for a gate to operate correctly, but they do not lead to a quantitative evaluation of its transfer characteristic. This can be done by considering the static noise immunity or *static noise margin* [45, 47].

Consider a chain of noninverting logic gates. Suppose that each gate is subject to noise such that the voltage at the input of the ith gate differs from the output voltage of the preceding gate, $v_{y(i-1)}$, by an amount

$$V_N = v_{xi} - v_{y(i-1)} \tag{5.1}$$

If sufficiently large, the noise voltage might cause the circuit output to go high in response to a low input level or vice versa—a condition that may quite appropriately be termed a *false change of state.* With a low-level input, a false change of state may occur if the noise voltage is positive; with a high-level input, this may occur when the noise voltage is negative.

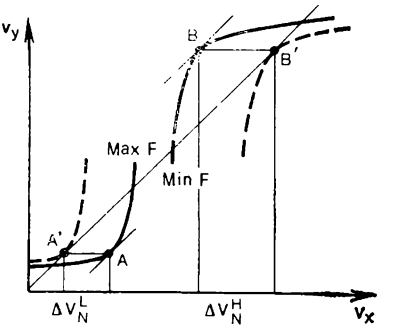


Fig. 5.10

In the former case, we shall define as the static noise margin, $\Delta V_N^{\mathbf{L}}$, the maximum level of positive noise voltage allowed at all the noninverting gates in cascade, that will not cause a false change of state. In the latter case, we shall define as the static noise margin, $\Delta V_N^{\mathbf{H}}$, the maximum level of the negative noise voltage allowed at all the noninverting gates in cascade, that will not cause a false change of state.

Graphically, the action of a noise voltage on a logic gate can be depicted as a horizontal shift of its transfer characteristic to the right when the noise voltage is negative and to the left when the noise voltage is positive. A false change of state will take place, if the shift removes points of intersection between the transfer characteristic and the line of unity gain. A gate will retain a noise margin, if the shift of the transfer characteristic is such that two adjacent points of intersection merge into one point of tangency. Then, as is seen from Fig. 5.10, the value of $\Delta V_N^{\mathbf{L}}$ will be given by the segment AA'. and the value of $\Delta V_N^{\mathbf{H}}$ by the segment BB'. The points A and B are points of unity gain, with point A lying on the Max F curve, and point B on the Min F curve. The static noise margin in the case of a logic 1 or logic 0 input can be expressed in terms of the coordinates of points A and B:

$$\Delta V_N^L = V_{Ax} - V_{Ay} \tag{5.2}$$

$$\Delta V_N^H = V_{By} - V_{Bx} \tag{5.3}$$

The noise immunity of inverting logic gates can be defined in a similar way, taking the transfer characteristic of a two-stage composite gate as a noninverting response. Configurations generating Max F are shown in Fig. 5.11, and Min F in Fig. 5.12.

Dynamic characteristics of logic gates. An important performance index of digital circuits and systems is their speed (or frequency response) defined as the number of operations carried out per unit time. An elementary operation (or a microoperation) usually consists in reading out a string of binary variables from storage, per-

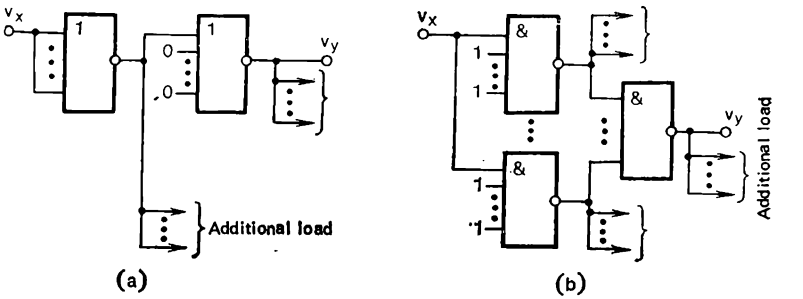


Fig. 5.11

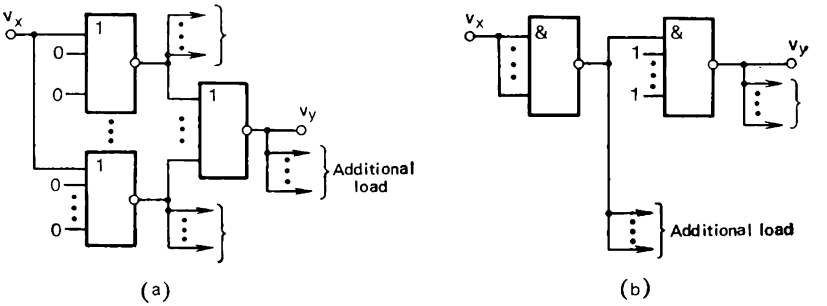


Fig. 5.12

forming the requested functional conversions, and reading the resultant switching functions back into storage. The execution of an elementary operation is initiated by a clock pulse applied to an appropriate input of the logic network. In circuits assembled from level logic gates, the clock pulse is the change of voltage which occurs at one of the inputs as the binary variable changes in value.

The execution of an operation may be visualized as the propagation of this change of voltage down the chain (or chains) of logic gates, with the last (final) gate serving usually as a part of storage. The time required to perform one elementary operation is determined by the propagation delay time defined as the time required for a gate to switch from its low output state to its high output state, or vice versa.

As is seen, the propagation delay time is a major dynamic characteristic of logic gates. Let τ_{pd}^{+} designate the propagation delay time when a gate goes from its low to its high output level, and τ_{pd}^{-}, when the gate goes the other way around. In the general case, $\tau_{pd}^{+} \neq \tau_{pd}^{-}$. For an inverting logic gate, we may also define the average propagation delay time, $\bar{\tau}_{pd} = (\tau_{pd}^{-} + \tau_{pd}^{+})/2$. Using the average propagation delay time, we can determine the overall or total propagation delay time for a chain formed by any even number of inverting logic gates.

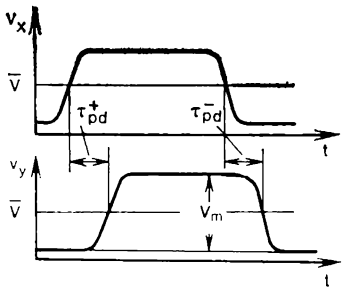


Fig. 5.13

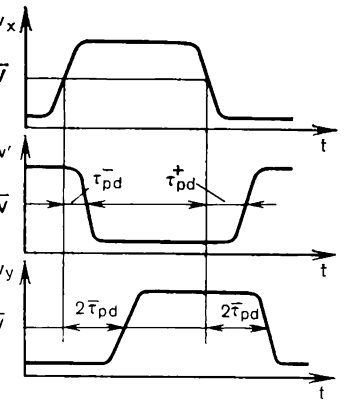


Fig. 5.14

In such a connection, the number of gates that go from their low to their high output state is equal to the number of gates that go from their high to their low output state. Therefore, for a number $2n$ of logic gates, the overall propagation delay time is

$$\Sigma\tau_{pd} = n\tau^-_{pd} + n\tau^+_{pd} = 2n\bar{\tau}_{pd}$$

With an arbitrary number n of gates in a chain, the average propagation delay time is approximately equal to the total propagation delay time, $\Sigma\tau_{pd} \approx n\bar{\tau}_{pd}$, but the uncertainty in the estimate will not increase with increasing n.

There is a scatter in the value of τ^+_{pd}, τ^-_{pd}, and $\bar{\tau}_{pd}$, which can be traced back to variations in manufacture and operating conditions. Most frequently, data sheets for logic gates quote the maximum value of the average propagation delay time, $\bar{\tau}_{pd,\max}$.

The manner in which the propagation delay time can be found from the timing diagrams of input and output voltages is illustrated in Fig. 5.13 for a noninverting logic gate and in Fig. 5.14 for an inverting logic gate. Since the output waveform usually follows the input waveform, any point on the waveform can be chosen as datum for reckoning the propagation delay time. In measurements, a convenient datum is provided by the point where the waveform crosses the $\bar{V}$ level, where $\bar{V} = (V^H + V^L)/2$ (see Figs. 5.13 and 5.14).

5.3. COMMON-EMITTER TRANSISTOR SWITCH LOGIC GATES

Inverter. An inverter (or a NOT gate) is basically a single common-emitter transistor switch (Fig. 5.15). The driving (control) circuit of the inverter consists of a base bias resistor R_B, a cutoff voltage

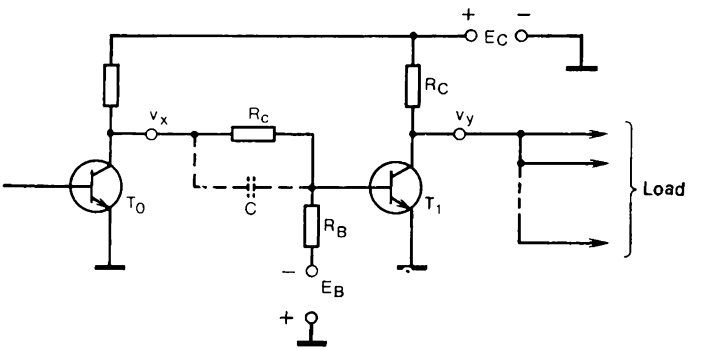


Fig. 5.15

source E_B, and a coupling resistor R_c which feeds the input signal to the base of T_1. To permit faster turn-on, the coupling resistor may be shunted by a speed-up capacitor, C, shown dashed in Fig. 5.15. It is assumed that the input voltage v_x is taken from the output of a similar device (transistor T_0 in Fig. 5.15).

An invertor may be built around transistors of any type of conduction. To make the matter more definite, we shall consider the circuit using *n-p-n* transistors widely used in microelectronics.

The transfer characteristic of an inverter, plotted for the simplified piecewise-linear transistor model (Fig. 5.16), has three segments, each representing a particular region (or condition) of transistor operation. Segment *1* applies when T_1 is at cutoff, and a high level voltage, V^H, exists at its collector. At no-load, V^H is a maximum:

$$V^H_{max} = E_C - I_{CBO}R_{CO} \approx E_C \tag{5.4}$$

Loading by the input resistance of a similar circuit leads to a decrease in output voltage. Because the OFF state of T_1 implies the ON (saturated) state of the transistors in the load circuit, the latter

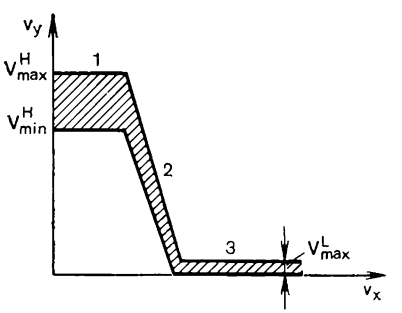


Fig. 5.16

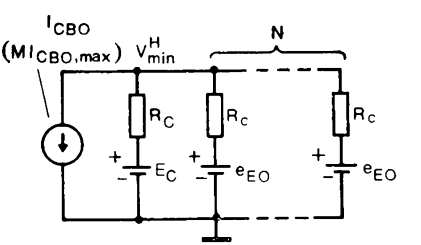


Fig. 5.17

may be visualized as consisting of a parallel combination of two active one-ports, R_c and the e_{EO} source (Fig. 5.17). With the fan-out equal to N, the lower limit of V^H can be found from Eq. (5.5)

$$V^H_{min} = V_{BE} + E_C \frac{\gamma_c}{N+\gamma_c}(1 - V_{BE}/E_C - I_{CBO,\,max}/I_{C(sat)}) \quad (5.5)$$

where $\gamma_c = R_c/R_C$ is the coupling factor and $I_{CBO,\,max}$ is the collector leakage current at maximum temperature.

Segment *2* applies when the transistor is in the active region where an approximately linear relation exists between the input and output voltages. At no-load, the gate has a maximum gain which decreases under load.

Segment *3* applies when the transistor is at saturation. The output voltage of the inverter is zero very nearly and, since the output resistance of a saturated transistor is low, it is practically independent of load. The upper limit for V^L is given by

$$V^L_{max} = V_{CE(sat)\,max} \quad (5.6)$$

where $V_{CE\,(sat)max}$ is the maximum collector-to-emitter voltage of the transistor in the saturated region.

Multi-input logic gates. A multi-input logic gate can be made with parallel-connected transistor switches operating into a common load (Fig. 5.18). The base circuits of the switches are analogous to that of the NOT gate (inverter) discussed above.

The output of the circuit in Fig. 5.18 goes high only when the transistors are all OFF at the same time, that is, when a low level voltage is applied to all the inputs simultaneously. In positive logic, this corresponds to the NOR switching function (the Pearce stroke).

When the parallel-connected transistors turn off all at the same time, the load resistor carries the sum of the collector leakage currents so that

$$V^H_{min} = V_{BE} + E_C \frac{\gamma_c}{N+\gamma_c} \times (1 - V_{BE}/E_C - MI_{CBO,\,max}/I_{C(sat)}) \quad (5.7)$$

Static noise margin. To find the static noise margin of a logic gate, it is legitimate to proceed from the following assumptions:

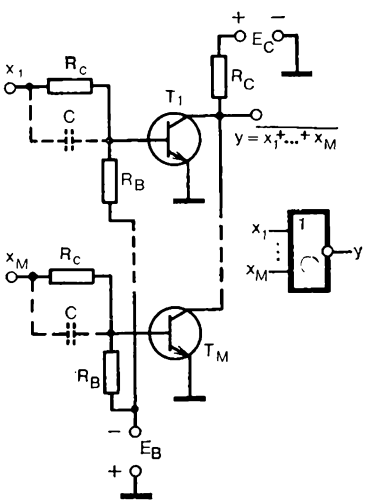


Fig. 5.18

1. The choice of ratings for R_C, R_c and R_B and of the supply voltages E_C and E_B is independent of the number of inputs the gate has and its loading. This assumption is applicable, when a logic gate is analyzed before a block diagram is developed for the associated digital circuit or system.

2. As regards the transistors used in a given gate, their parameters can be improved without bound. That is, it is assumed that $\beta_{max} = \infty$, $I_{CBO,\,min} = 0$, and $V_{CE(sat)min} = 0$. This assumption simplifies the design relations at the cost of an insignificant derating of ΔV_N^L and ΔV_N^H.

3. The change in the base voltage of a conducting transistor is negligible in comparison with V_{BE} of a saturated transistor. That is, it is legitimate to assume that the cut-in voltage of a transistor is equal to its voltage at saturation:

$$V_{BE(sat)} \approx e_{EO}$$

Since the base voltage is assumed constant, we may define E_B and R_B in terms of one quantity, namely the bias current I_{bias} flowing through R_B when the transistor is conducting:

$$I_{bias} = (E_B + e_{EO})/R_B \tag{5.8}$$

Static noise immunity is calculated using the procedure set forth in Sec. 5.2 and consists in locating the points A and B on the transfer characteristic of the composite two-stage logic gate (Fig. 5.19). In a piecewise-linear approximation of the transfer characteristic, the coordinates along the v_y axis are defined by logic levels 0 and 1:

$$V_{Ay} = V_{max}^L, \qquad V_{By} = V_{min}^H \tag{5.9}$$

The coordinate of point A on the v_x axis can be found by inspection of the logic gates connected in cascade (Fig. 5.20). As already noted, for Max F to be generated, v_y must be maximized and V' must at the same time be minimized. For the type of logic gate in question, the effect of loading manifests itself in bringing down the output voltage, so in the circuit of Fig. 5.20 the additional load is

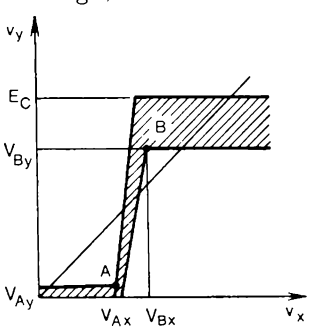


Fig. 5.19

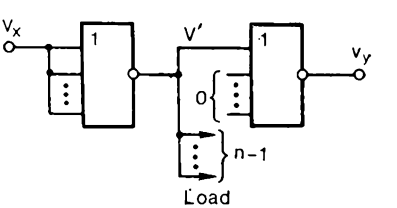


Fig. 5.20

shown connected to the output of only the first stage. In addition to loading, the output voltage of a logic gate is affected by the values of β and I_{CBO}. As they are raised, the collector voltage builds up, and this leads to a decrease in output voltage. Therefore, we require that the transistors in the first stage have maximum values of β and I_{CBO}, and those of the second stage, minimum values of β and I_{CBO}, that is,

$$\beta_1 = \infty, \quad I_{CBO,\,1} = I_{CBO,\,max}, \quad \beta_2 = \beta_{min}, \quad I_{CBO,\,2} = 0$$

Since at point A the first-stage transistor is in the active region, its base current is

$$i_{B1} = (I_{C1} - I_{CBO})/\beta_1 - I_{CBO} = -I_{CBO,\,max} \tag{5.10}$$

Using the equivalent base circuit for the first transistor (Fig. 5.21), we find v_x at point A to be:

$$V_{Ax} = e_{EO} + (I_{bias} - I_{CBO,\,max})\,\gamma_c R_C \tag{5.11}$$

The coordinate of point B on the v_y axis is found by inspection of the logic gate in cascade in Fig. 5.22. In contrast to the previous case, for Min F to be generated it is essential to minimize v_y and, at the same time, maximize V'. To this end, the second stage is loaded as much as possible, and the transistor parameters are required to be as follows: $\beta_1 = \beta_{min}$, $I_{CBO} = 0$, $\beta_2 = \infty$, $I_{CBO,2} = I_{CBO,max}$.

At point B, the second-stage transistors are at the cut-in boundary, so

$$V_{BE,\,2} = e_{EO}, \quad I_{B2} = -I_{CBO,\,max} \tag{5.12}$$

Using the equivalent circuit for interstage coupling in Fig. 5.23, we can calculate V' and I_{C1} corresponding to point B:

$$\begin{aligned} V' &= e_{EO} + (I_{bias} - I_{CBO,\,max})\,R_C\gamma_c \\ I_{C1} &= (E_C/R_C)\,(1 - e_{EO}/E_C) - (1+\gamma_c)\,(I_{bias} - I_{CBO,\,max}) \end{aligned} \tag{5.13}$$

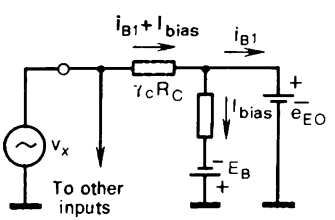


Fig. 5.21

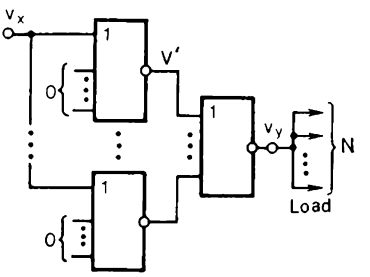


Fig. 5.22

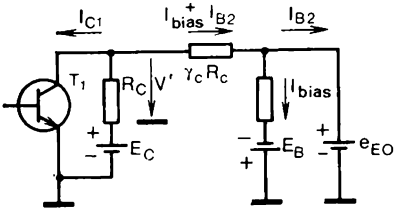


Fig. 5.23

The value of v_x corresponding to point B can be found, using the equivalent circuit of Fig. 5.21, assuming that $I_{B1} = I_{C1}/\beta_{min}$:

$$\begin{aligned} V_{Bx} &= e_{EO} + (I_{bias} + I_{B1})\,\gamma_c R_C \\ &= e_{EO} + (\gamma_c E_C/\beta_{min})\,(1 - e_{EO}/E_C \\ &\quad + \beta_{min} I_{CBO,\,max}/I_{C(sat)}) \\ &\quad + \gamma_c R_C\,(I_{bias} - I_{CBO,\,max}) \\ &\quad \times [1 - (1+\gamma_c)/\beta_{min}] \end{aligned} \tag{5.14}$$

Substituting Eqs. (5.6), (5.7), (5.11), and (5.14) into Eqs. (5.2) and (5.3) yields an expression for the noise margin

$$\Delta V_N^L = e_{EO} - I_{CE(sat)} + (I_{bias} - I_{CBO,\,max})\,\gamma_c R_C \tag{5.15}$$

$$\begin{aligned} \Delta V_N^H &= \gamma_c E_C \left(\frac{c_1}{N+\gamma_c} - \frac{c_2}{\beta_{min}}\right) - \left(1 - \frac{1+\gamma_c}{\beta_{min}}\right) \\ &\quad \times (I_{bias} - I_{CBO,\,max})\,\gamma_c R_C \end{aligned} \tag{5.16}$$

where

$$c_1 = 1 - e_{EO}/E_C - M I_{CBO,\,max}/I_{C(sat)} \tag{5.17}$$

$$c_2 = 1 - e_{EO}/E_C + \beta_{min} I_{CBO,\,max}/I_{C(sat)} \tag{5.18}$$

Equations (5.15) and (5.16) indicate that the noise immunity of a logic gate substantially depends on I_{bias} and γ_c and decreases with increasing fan-in and fan-out.

Operability conditions. The operability of a logic gate is ensured within the range of circuit parameter values where ΔV_N^L and ΔV_N^H are positive. Physically, the operability of a logic gate depends on the same set of factors that are responsible for the transistors used in the gate to be driven to cutoff and saturation. In electric calculations, the values of R_C, E_C, and E_B are usually found for the specified power consumption and speed (frequency response). The operability is ensured through a proper choice of values for R_c and R_B. Since the values of R_c and R_B can unambiguously be expressed in terms of γ_c and I_{bias} [$R_c = R_C\gamma_c$ and $R_B = (E_B + e_{EO})/I_{bias}$], it is convenient to state the operability conditions for a logic gate as the constraints imposed on the values of γ_c and I_{bias}.

Solving the inequalities $\Delta V_N^{\mathrm{L}} > 0$ and $\Delta V_N^{\mathrm{H}} > 0$ simultaneously for I_{bias} gives the operability conditions as

$$I_{\mathrm{bias}} > I_{\mathrm{bias,\ min}} = I_{\mathrm{CBO,\ max}} - \frac{e_{\mathrm{EO}} - V_{\mathrm{CE(sat)}}}{\gamma_{\mathrm{c}} R_{\mathrm{C}}} \tag{5.19}$$

$$I_{\mathrm{bias}} < I_{\mathrm{bias,\ max}} = I_{\mathrm{CBO,\ max}} + \frac{I_{\mathrm{C(sat)}}}{\beta_{\mathrm{min}} - 1 - \gamma_{\mathrm{c}}} \left(\frac{c_1 \beta_{\mathrm{min}}}{N + \gamma_{\mathrm{c}}} - c_2 \right) \tag{5.20}$$

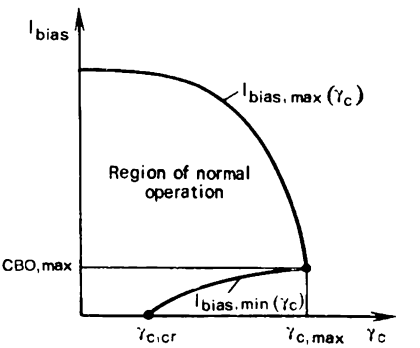


Fig. 5.24

Inequality (5.19) corresponds to the condition for the transistor to be driven to cutoff by a low level applied to the input of the logic gate, and inequality (5.20) to the condition for the transistor to be driven into the saturation region by a high level applied to the gate's input.

On the I_{bias} v.s. γ_{c} diagram (Fig. 5.24), the region within which a logic gate is operable is bounded by the $I_{\mathrm{bias,min}}(\gamma_{\mathrm{c}})$ and $I_{\mathrm{bias,\ max}}(\gamma_{\mathrm{c}})$ curves. The intersection of the curves locates the limiting value of γ_{c}. Assuming that at the intersection $I_{\mathrm{bias,min}} \approx I_{\mathrm{CBO,max}}$, we obtain

$$\gamma_{\mathrm{c,\ max}} \approx (c_1/c_2)\,\beta_{\mathrm{min}} - N \tag{5.21}$$

Within the region where $\gamma_{\mathrm{c}} < \gamma_{\mathrm{c,cr}}$, the condition for operability is satisfied at $I_{\mathrm{bias}} = 0$. The value of $\gamma_{\mathrm{c,cr}}$ can be found from Eq. (5.19) at $I_{\mathrm{bias,min}} = 0$:

$$\gamma_{\mathrm{c,\ cr}} = \frac{e_{\mathrm{EO}} - V_{\mathrm{CE(sat)}}}{R_{\mathrm{C}} I_{\mathrm{CBO,\ max}}} \tag{5.22}$$

Optimization of static noise immunity. The operability conditions of a logic gate define the range of limiting values for I_{bias} and γ_{c}. Therefore, I_{bias} and γ_{c} must be chosen such that the noise immunity of the gate is a maximum.

An analysis of Eqs. (5.15) and (5.16) will show that an increase in I_{bias} entails an accompanying increase in ΔV_N^{L} and an accompanying decrease in ΔV_N^{H}, and an optimal noise margin is attained at a bias current such that $\Delta V_N^{\mathrm{L}} = \Delta V_N^{\mathrm{H}}$. Solving this equation for I_{bias} yields

$$I_{\mathrm{bias,\ opt}} = I_{\mathrm{CBO,\ max}} + \frac{\beta_{\mathrm{min}}}{2\beta_{\mathrm{min}} - 1 - \gamma_{\mathrm{c}}} \times \left[V_{\mathrm{CE(sat)}} - e_{\mathrm{EO}} + \gamma_{\mathrm{c}} E_{\mathrm{C}} \left(\frac{c_1}{N + \gamma_{\mathrm{c}}} - \frac{c_2}{\beta_{\mathrm{min}}} \right) \right] \frac{1}{\gamma_{\mathrm{c}} R_{\mathrm{C}}} \tag{5.23}$$

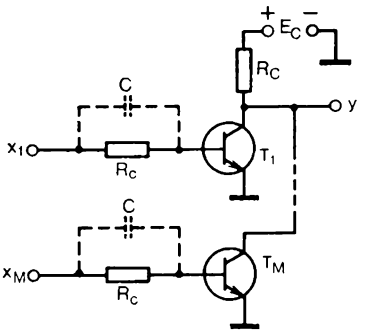


Fig. 5.25

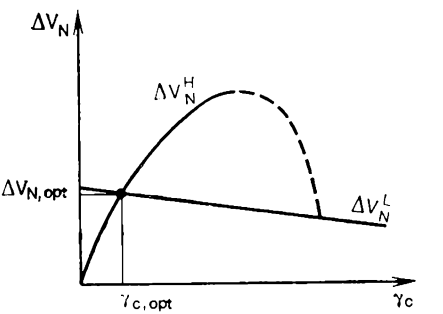


Fig. 5.26

On substituting Eq. (5.23) in Eq. (5.15), we can find the noise margin ΔV_N for an optimal bias current. In view of the equality $\Delta V_N^H = \Delta V_N^L$ we have

$$\Delta V_N = e_{EO} - V_{CE(sat)} + \frac{\beta_{min}}{2\beta_{min} - 1 - \gamma_c} \times \left[V_{CE(sat)} - e_{EO} + \gamma_c E_C \left(\frac{c_1}{N + \gamma_c} - \frac{c_2}{\beta_{min}} \right) \right] \quad (5.24)$$

Let us choose $\gamma_c = \gamma_{c,opt}$ such that ΔV_N is a maximum. To find the maximum of ΔV_N (γ_c), we evaluate the derivative d $(\Delta V_N)/d\gamma_c$, equate it to zero, and find $\gamma_{c,opt}$. Dropping the terms which are negligible in comparison with unity and setting $c_1 \approx c_2 \approx 1$, we get

$$\gamma_{c,\,opt} \approx 2N \left(\sqrt{1 + \beta_{min}/2N} - 1 \right) \quad (5.25)$$

Unbiased logic gates. As we have found, a logic gate is able to operate at zero bias current. Therefore, we can safely simplify its circuit by dropping R_B and E_B (Fig. 5.25). Despite the reduced noise immunity, unbiased logic gates are widely used in resistor-transistor logic (RTL) and resistor-capacitor-transistor logic (RCTL) gates. Consider the calculation and optimization of noise immunity for such gates.

On setting $I_{bias} = 0$ in Eqs. (5.15) and (5.16), we obtain expressions for the noise margin of RTL gates:

$$\Delta V_N^L = e_{EO} - V_{CE(sat)} - I_{CBO,\,max} \gamma_c R_C \quad (5.26)$$

$$\Delta V_N^H = \gamma_c E_C \left(\frac{c_1}{N + \gamma_c} - \frac{c_2}{\beta_{min}} \right) + \left(1 - \frac{1 + \gamma_c}{\beta_{min}} \right) I_{CBO,\,max} \gamma_c R_C$$

Plots of ΔV_N^L and ΔV_N^H as functions of γ_c appear in Fig. 5.26. The noise margin is optimal at point a where

$$\Delta V_N^L = \Delta V_N^H \quad (5.27)$$

Solving Eq. (5.27) and dropping infinitesimal terms gives

$$\gamma_{c,\,opt} \approx N\,(e_{EO} - V_{CE(sat)})/E_C\,(1 - 2e_{EO}/E_C - N/\beta_{min}) \quad (5.28)$$

Eq. (5.28) shows that high values of fan-out, N, call for higher R_c.

Power dissipation. The total power consumption P is the sum of the powers drawn by a logic gate from the collector supply, P_C, and the base supply, P_B. Practically, $E_C \gg E_B$ nearly always, and I_{R_C} flowing through R_C is many times I_{bias}. Therefore, P_B can be neglected in comparison with P_C, so that

$$P \approx P_C = EI_{R_C} = E\,(E - v_y)/R_C$$

In the logic 0 state, $v_y = V_{CE(sat)} \approx 0$, so $P^0 E_C^2/R_C$.
In the logic 1 state, $v_y = V^H$. Using Eq. (5.10) and neglecting the infinitesimal terms gives

$$P^1 \approx (E_C^2/R_C)\,N/(N + \gamma_c)$$

The average power is given by

$$\bar{P} \approx (E_C^2/2R_C)\left(1 + \frac{N}{N + \gamma_c}\right) \quad (5.29)$$

On substituting Eq. (5.25) or (5.28) for γ_c in Eq. (5.29), we can obtain the average power for optimized circuits.

Transient Response

Consider the transients in a cascade of three common-emitter transistor switch logic gates (Fig. 5.27), assuming that

—each gate operates at a maximum fan-in and fan-out, M_{max} and N_{max};

—in each gate, only one transistor is switched at a time, with the other (M — 1) transistors remaining in the OFF state;

—in all gates of the cascade, the transients accompanying a transition from a zero to a 1 state or back repeat one another except that they are shifted in time;

—the change of voltage at the output of each gate is a maximum, that is, $V^L = V^L_{min} = 0$ and $V^H = V^H_{max}$.

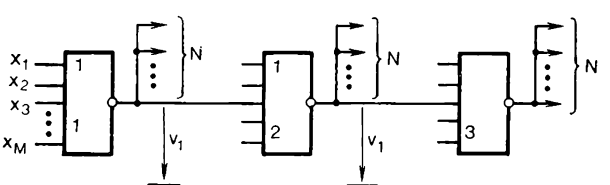


Fig. 5.27

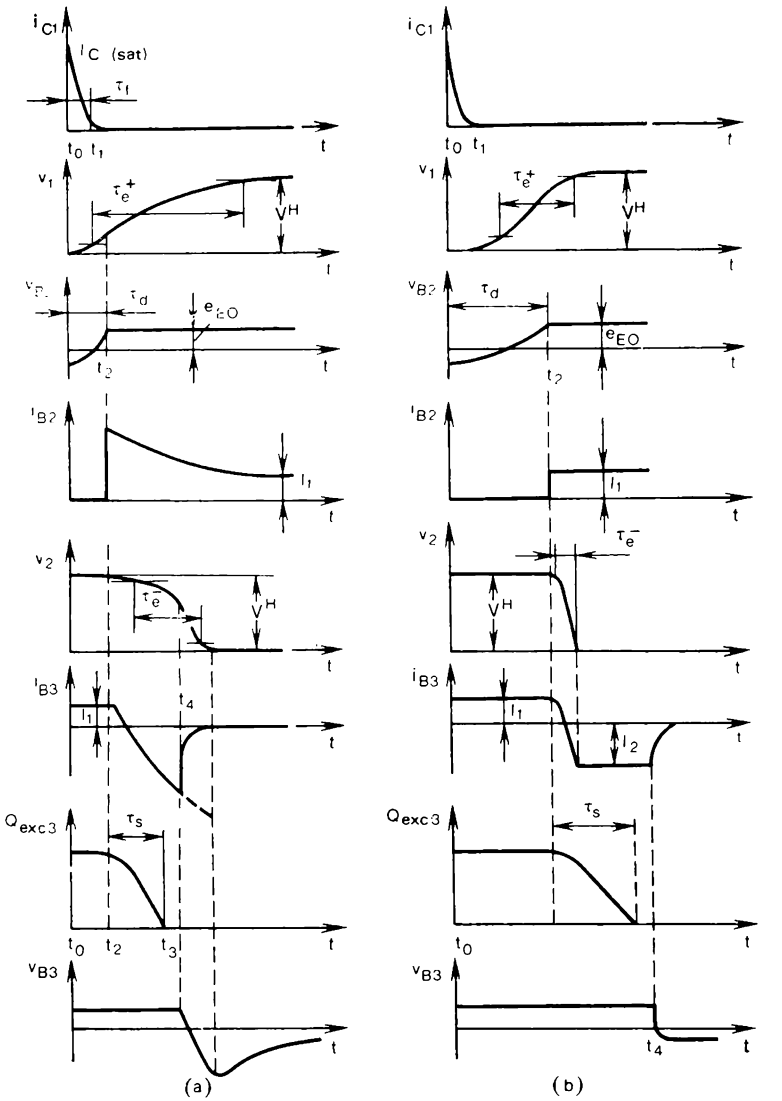


Fig. 5.28

Let us trace the response of the cascade in two cases, namely for $C > 0$ (Fig. 5.28*a*) and $C = 0$ (Fig. 5.28*b*). Transistor $T2$ begins to change state at time t_0 when transistor $T1$ moves from the saturated to the active region and the decrease in the collector current, i_{C1}, re-charges the stray circuit capacitances, with the result that

v_1 and v_{B_2} gradually build up. At time t_2, v_{B_2} rises to the cut-in value for $T2$, after which v_{B_2} remains practically unaltered because of the low input resistance of the conducting transistor. Transistor $T1$ usually turns off before $T2$ turns on (at time t_1). In a circuit using a speed-up capacitor and low-frequency alloyed transistors it may so happen that $t_1 > t_2$.

Beginning with t_2, the collector current of $T2$ builds up as the excess minority-carrier charge is being removed from $T3$. At time t_3, the excess minority-carrier charge Q_{exc} reduces to zero, and $T3$ moves into the active region. Because time t_3 at which $T3$ is switched corresponds to time t_0 when $T1$ is switched, the time interval from t_0 to t_3 is equal to the propagation delay time through $T2$ and $T3$.

In a circuit without a speed-up capacitor, the leading and trailing edge times for the positive- and negative-going pulses, τ_e^+ and τ_e^-, are usually shorter than the delay time, τ_d and the storage time, τ_s. The speed-up capacitor markedly cuts down the propagation delay time. Unfortunately, the time required to charge it up is added to τ_e^+ and a negative voltage spike is developed at the base of the transistor being turned off; this spike is sometimes called the dynamic bias of a transistor.

The technique and procedure for evaluating the delay, rise, storage, and fall times of the collector current have been examined in sufficient detail in Chap. 3.

5.4. MULTISTEP LOGIC GATES

In the common-emitter transistor switch logic gates examined so far, the coupling resistors serve to match the input and output signals in terms of parameters. In contrast, the coupling circuits (usually, nonlinear ones) in multistep logic gates contribute to the implementation of a switching function, thereby permitting an increase in the number of inputs per inverting transistor. The nonlinear elements most commonly used in these coupling circuits are diodes (diode-transistor logic, or DTL, gates) and transistors (transistor-transistor logic, or TTL, gates).

DTL Gates

The circuit of a DTL gate is shown in Fig. 5.29. When all the M inputs are high, the input diodes D_1, D_2, . . ., D_M are reverse-biased, and the transistor base draws a heavy current equal to the difference of currents through R_a and R_B (the small currents flowing through the reverse-biased diodes may be neglected). The transistor is saturated, and a low voltage level (close to zero) exists at its collector: $V^L = V_{CE(sat)} \approx 0$. For the transistor to be turned off it

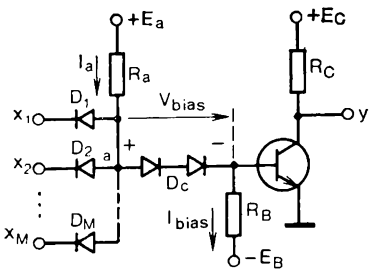


Fig. 5.29

is essential that a low logic level, V^{L}, should exist at least at one input. Then the respective input diode will be forward-biased, and it will pass a major proportion of the current flowing through R_a. As a result, a high voltage level close to the collector supply voltage E_C will exist at its collector: $V^{H} \approx E_C$. In the case of positive logic, this manner of generating a low and a high voltage level by a DTL gate is described by the NAND switching function:

$$y = \overline{x_1, x_2, \ldots, x_M}$$

As already noted, interstage coupling is provided by a series chain of coupling diodes, D_c. Owing to the R_B-E_B bias circuit the operating point of the coupling diodes is shifted into a high, slope portion of the V-I characteristic. The difference between the maximum and minimum bias voltages in the chain of coupling diodes is small, so that in practice it may safely be assumed that

$$V_{\text{bias, max}} \approx V_{\text{bias, min}} = 2V_{f,\,D_c} \approx 1.4 \text{ to } 1.6 \text{ V}$$

Consider the static operating conditions required for transistor to move into the cutoff and saturated regions. The change in the potential at point a occurring as a result of transistor switching gives rise to a change in the currents flowing through R_a and R_B. When the transistor is at cutoff, let these currents be I_a^0 and I_{bias}^0, and at saturation I_a^1 and I_{bias}^2. When the transistor is at cutoff, the input signal will forward-bias at least one of the input diodes. Therefore, the potential at point a is equal to the sum

$$V^{L} + V_{f,\,D_{\text{in}}} = V_{\text{CE(sat)}} + V_{f,\,D_{\text{in}}} \approx 0.8 \text{ to } 1.0 \text{ V}$$

When the transistor is at saturation, the potential at point a is equal to the sum

$$e_{EO} + 2V_{f,\,D_{\text{in}}} \approx 1.8 \text{ to } 2.2 \text{ V}$$

In the circumstances,

$$I_a^0 = (E_a - V_{\text{CE(sat)}} - V_{f,\,D_{\text{in}}})/R_a$$

$$I_a^1 = (E_a - 2V_{f,\,D_c} - e_{EO})/R_a$$

The base potential is below the potential at point a by $V_{\text{bias}} = 2V_{f,\,D_c}$. Therefore,

$$I_{\text{bias}}^0 = (E_B + V_{\text{CE(sat)}} + V_{f,\,D_{\text{in}}} - 2V_{f,\,D_c})/R_B$$

$$I_{\text{bias}}^1 = (E_B + e_{EO})/R_B$$

With silicon transistors, it is customary to drop the bias source, E_B, and to return R_B directly to the common ground. Then, $I^0_{bias} \approx 0$.

For the transistor to move into saturation, its base and collector currents must be such that $i_B\beta_{min} > I_{C(sat)}$. The base current of a saturated transistor is $I^1_a - I^1_{bias}$; its collector current depends on load. Consider the worst case where the load current is a maximum: $I_L = NI^0_a$; then the condition for saturation may be written

$$N < [\beta_{min}(I^1_{a,\,min} - I^1_{bias,\,max}) - I_{R_C}]/I^0_{a,\,max} \qquad (5.30)$$

Inequality (5.30) defines the maximum fan-out for a logic gate. Among other things, it follows that in order to improve fan-out (to increase the value of N), it is desirable to minimize the relative change in I_a as the transistor is switched, because the ratio I^1_a/I^0_a will then increase. To achieve this goal, the value of E_a is chosen to be several times the change of potential at point a; as a rule, $E_a = 5$ to 7 V. Any further increase in E_a is not warranted, as it would lead to an increased power drain.

In a biased circuit (with a source of E_B included), the transistor is driven to cutoff when at least one of the inputs is low. This happens because V_{bias} exceeds in absolute value the potential at point a, and the base goes negative. In an unbiased circuit (containing no E_B supply), the voltage level at point a is not sufficient to shift the operating point of the coupling diodes into the high-slope portion of their V-I characteristic, and the total resistance of the coupling diodes turns out considerably higher than R_B, so the base voltage is zero very nearly.

Noise immunity. For noise immunity analysis, we may use the approach set forth in connection with resistance-coupled gates. The noise margin for a low-level input (defining quantitatively the reliability of transistor turn-off) can be found from an equation similar to Eq. (5.15)

$$\Delta V^L_N = e_{EO} - V_{CE(sat)} + 2V_{f,\,D_c} - V_{f,\,D_{in}} \qquad (5.31)$$

The difference term, $2V_{f,\,D_c} - V_{f,\,D_{in}}$, gives the total voltage drop across D_{in} and D_c between the input of the logic gate and the transistor base and replaces the term $(I_{bias} - I_{CBO,max})\,R_c$ in Eq. (5.15). The currents flowing through D_c and D_{in} when the base current of the first logic gate is zero (the worst case) are:

$$I_{D,\,c} \approx I^1_{bias}, \quad i_{D,\,in} \approx I^1_a - I^1_{bias}$$

On the one hand, an increase in I^1_{bias} leads to an increase in the difference term, $2V_{f,\,D_c} - V_{f,\,D_{in}}$, and, as a consequence, to an increase in ΔV^L_N. On the other hand, this leads to a decrease in the saturation coefficient of the transistor, $K_{sat} = (I^1_a - I^1_{bias})\,\beta/(I_{R_C}$

$+ NI_a^0$), which entails an increase in $V_{CE(sat)}$ and a decrease in ΔV_N^L. There is an optimal value for I_{bias}^1 at which ΔV_N^L is a maximum. Using Eqs. (3.4) and (3.24) to evaluate the quantities entering Eq. (5.31), we can reduce the optimization of the logic gate to finding a maximum for the term

$$[I_a^1(1-N/\beta-I_{R_C}/\beta I_a^1)-I_{bias}^1](I_{bias}^1)^2/(I_a^1-I_{bias}^1)$$

The sought optimum is noncritical because the voltage across the p-n junction is a function of current. On the other hand, the likely spread in parameters between the circuit components might result in a heavy bias current, thereby upsetting the condition for the transistor to move into the saturation region and impairing the operability of the logic gate. Therefore, it is advisable to seek an optimal bias current for β_{min}, $I_{a,min}^1$, and $I_{R_C,max}$ because these extremal values correspond to the worst conditions for a transistor to move into saturation. Then

$$I_{bias,\,opt}^1 \approx 0.66 I_{a,\,min}^1 (1-N/\beta_{min}-I_{R_C max}/\beta_{min} I_{a,\,min}^1) \qquad (5.32)$$

At the optimal bias current, the typical value of ΔV_N^L is from 1 to 1.2 V.

The noise margin with a high-level input is approximately equal to the reverse voltage across the input diode, (see piecewise-linear model of Fig. 5.7) provided the voltage applied to the input of the logic gate is $V^H \approx E_C$. That is,

$$\Delta V_N^H \approx E_C - V_a^1 + e_{OD_{in}} = E_C - e_{EO} - 2V_{f,\,D_c} + e_{OD_{in}} \qquad (5.33)$$

In practice, E_C is taken to be anywhere between 2.5 and 3 V, so that ΔV_N^H is close to ΔV_N^L. Any further increase in E_C is unwarranted, as it would lead to an increased power drain and reduced speed of the logic gate.

Power dissipation. In the general case, the total power dissipation is a sum of two terms: $E_a I_a + E_C I_{R_C}$. Substituting the values of currents corresponding to the two logic states of the gate, we obtain

$$P^0 = E_a(E_a - 2V_{f,\,D_c} - e_{EO})/R_a + E_C^2/R_C \qquad (5.34)$$

$$P^1 = E_a(E_a - V_{f,\,D_{in}} - V_{CE(sat)})/R_a \qquad (5.35)$$

$$\bar{P} \approx E_a(E_a - 1.2\text{V})/R_a + E_C^2/2R_C \qquad (5.36)$$

As follows from the above relations, the power dissipation is a function of R_a and R_C. For state-of-the-art DTL gates, $\bar{P}$ ranges from 10 to 20 mW.

Transient response. Consider the transient response of a chain of three identical DTL gates connected as shown in Fig. 5.27. to

an input signal. As $LG1$ turns off, $LG2$ turns on with a delay, τ_d, equal to the time interval from the instant when $LG1$ moves out of saturation to the instant when $LG2$ turns on. The circuit parameters are usually chosen such that the initial rate of rise for the output voltage v_1 of $LG1$ is higher than that for the voltage at point a of $LG2$, that is

$$I_{R_C}/C_{\text{out}} > (I_a^0 - I_{\text{bias}}^0)/(C_a + C_{\text{in}}) \tag{5.37}$$

where

$$C_{\text{out}} = C_c + C_{\text{c-s}}^3 + C_{R,\text{C}} + (N-1)\,C_{D,\text{in}} + C_{\text{w}}$$

$$C_a = (M-1)\,C_{D,\text{in}} + C_{\text{C-s}}\,(M+2) + C_{R,a}$$

$$C_{\text{in}} = C_e + C_c$$

C_{out} = equivalent capacitance at the output of $LG1$
C_a = equivalent capacitance at point a
C_{in} = input capacitance of the transistor
$C_{R,a}$ = capacitance of R_a
$C_{R,\text{C}}$ = capacitance of R_{C}
$C_{D,\text{in}}$ = capacitance of the input diode
$C_{\text{C-s}}$ = collector-to-substrate capacitance
C_{w} = wiring (circuit) capacitance

When the condition defined by Eq. (5.37) is satisfied, the input diode is driven to cutoff very rapidly, so the voltage at point a rises due to the action of the difference current $I_a - I_{\text{bias}}$ the average value of which over the delay time, τ_d, is $(I_a^0 + I_a^1 - I_{\text{bias}}^0 - I_{\text{bias}}^1)/2$ Hence,

$$\tau_d \approx 2\,(C_a + C_{in})\,\Delta v_a/(I_a^0 + I_a^1 - I_{\text{bias}}^0 - I_{\text{bias}}^1)$$

where Δv_a is the change of voltage at point a over the interval τ_d, approximately equal to the difference between the high and low levels (about 1.2 V).

The time interval between the instant when the transistor of $LG2$ turns on and the instant when the transistor of $LG3$ moves out of saturation gives the storage time, τ_s. This interval is the sum of the rise time, τ_e^-, for the negative change of voltage at the collector of $T2$ and the time required for the removal of the minority-carrier saturation charge stored in the base. Both the negative change of voltage and the removal of the excess charge proceed at a constant base current and may be evaluated, using Eqs. (3.47) and (3.52). For $T2$, the cut-in (turn-on) current is $I_1 = I_a^1 - I_{\text{bias}}^1$. In determining the turn-off current for $T3$ it is essential to include the reverse current flowing through the coupling diodes due to the removal of the minority-carrier saturation charge stored in the bases of the diodes. If the diode charge exceeds the excess minority-carrier charge of the saturated transistor, the coupling diodes will be conducting over the entire turn-off delay time of $T3$, and the turn-off cur-

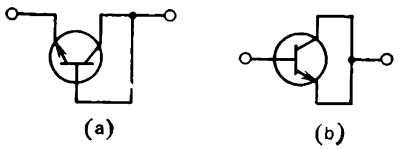


Fig. 5.30

ent will be a function of the input resistance of the transistor:

$$I_2 \approx (V_{\text{bias}} + V_{\text{BE}} - V_{\text{CE(sat)}} - V_{f,\,D_{\text{in}}})/r_{\text{in}}$$

The accelerated turn-off of the transistor aids in stepping up the speed of a DTL gate in comparison with a RTL gate.

Omission of capacitors and high-value resistors makes DTL gates convenient for use in IC form. In ICs, the function of diodes is performed by transistors connected as diodes. The input diode is connected as shown in Fig. 5.30*a*. The connection of a transistor as a diode in the manner shown in Fig. 5.30*b* ensures that in the ON state a minority-carrier charge is stored in both the base and the collector. Owing to this arrangement, the reverse recovery time of the coupling diode is sufficient for an accelerated turn-off of the transistor.

TTL Gates

Gates in this family are very close to DTL gates in both the operating principle and basic behaviour [24, 26]. A major distinction is that the diode structure at the input of the gate is replaced by a *multi-emitter transistor*, as shown in Fig. 5.31. The emitter junctions of the multi-emitter transistor T_c act as input diodes, and the collector junction as a bias diode. An advantage of a TTL gate over a DTL gate is that it is better adapted to integrated-circuit technology. A multi-emitter transistor takes up a smaller area on the chip in comparison with a diode structure. Also, owing to the manner in which the transistor operates at the gate's input, it is possible to drop R_B which ordinarily takes up a considerable area. At present, TTL gates are the most widely used family of integrated circuits. In discrete form, they are not used.

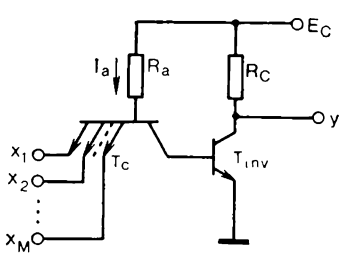


Fig. 5.31

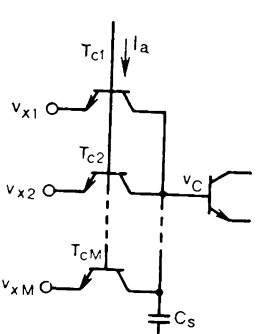


Fig. 5.32

A multi-emitter transistor may be treated as a combination of several n-p-n transistors (Fig. 5.32), with a stray insulation capacitance, C_s, connected to their common collector.

A high level applied simultaneously to all the inputs of the gate reverse-biases the emitter junctions of transistors $T_{c1}, \ldots, T_{c,M}$, so that they move into the inverted active region. Now, the input current is governed by the inverse current gain, β_I, of transistors $T_{c1}, \ldots, T_{c,M}$, namely:

$$I^{H}_{in,\,i} = I_a \beta_{I,\,i} \quad (i = 1, 2, \ldots, M)$$

whence it follows that in order to reduce the input currents, it is necessary to reduce the value of β_I.

When a low level exists at one or several inputs, the respective transistors move into the saturation region. I^{L}_{in} is a maximum when only one of the transistors in the multi-emitter structure is saturated. Since the collector current of a multi-emitter transistor is zero very nearly, the input current of the logic gate is given by

$$I^{L}_{in,\,max} \approx I_a + (M-1)\, I^{H}_{in,\,i}$$

Now, the saturation voltage of the multi-emitter transistor is

$$V^{*}_{CE(sat)} = m\varphi_T \ln [(1 + M\beta_I)/(M'\beta_I)]$$

where M' is the number of emitters going low at a given instant of time.

The decrease in β_I owing to a rise in $V_{CE(sat)}$ and, as a consequence, an increase in the base voltage of the inverting transistor leads to a reduced static noise margin for the logic gate in the case of a low-level input. Therefore, β_I ought never to be chosen less than $0.05/M$. Then $V^{*}_{CE(sat)}$ will not exceed 120 mV. Generally transients in TTL gates are similar to those in DTL gates.

TTL Gates with a Totem-Pole Output Driver

The fan-in, fan-out, noise margin and speed (frequency response) of a multi-step logic gate can be enhanced by adding a push-pull final transistor stage to the main inverting transistor. This output configuration is known as a totem-pole amplifier. It is also referred to as a power-driver output stage. The most commonly used circuit, a TTL NAND gate, is shown in Fig. 5.33. DTL gates are built along similar lines.

Consider the static transfer characteristic of the TTL NAND gate, using piecewise-linear approximation (Fig. 5.34). When at least one of the inputs is low, the respective n-p-n structure of the multi-emitter transistor is saturated, and the low voltage existing at the

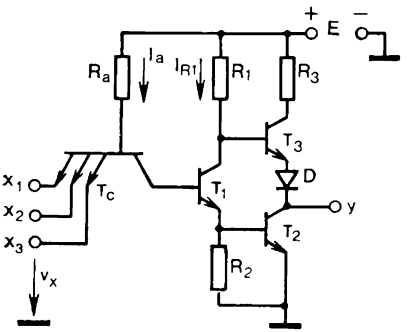


Fig. 5.33

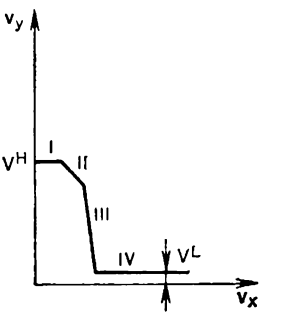


Fig. 5.34

base of $T1$ is not sufficient to drive it into conduction. Transistors $T1$ and $T2$ are OFF, whereas transistor $T3$ and the diode D are conducting and maintain a high output level, V^H (portion I) despite the low output resistance.

As the input voltage rises, transistor $T1$ turns on, and the operating point shifts into portion II of the transfer characteristic, where the slope is defined by the ratio R_1/R_2. The further rise in the input voltage causes transistor $T2$ to turn on, and the accompanying increase in gain increases the slope of the transfer characteristic (portion III).

After $T2$ moves into the saturated region, the output of the logic gate goes low and remains at a practically constant voltage, V^L (portion IV). When all inputs of the logic gate are high, the emitter junctions of T_c are reverse-biased, and I_a is fully diverted into the base of $T1$ so that it finds itself saturated. $T3$ and D are OFF, because the difference of potentials between the collectors of $T1$ and $T2$ is not sufficient to forward-bias the two series-connected p-n junctions.

The fan-in and fan-out of a TTL gate with a totem-pole output driver is increased in comparison with the basic TTL gate of Fig. 5.31. The point is that the output transistor $T2$ goes to saturation more readily because the base current is incremented by I_{R1} and the collector current is decremented by $I_{R,C}$ when the gate output goes low, and also because the output resistance decreases when the gate output goes high.

Immunity towards a positive noise voltage is improved, because at the instant when $T2$ turns on the input voltage is governed by two emitter junctions. Therefore, the noise margin is

$$\Delta V_N^L \approx 2e_{EO} - V_{CE(sat)} - V_{CE(sat)}^* \approx 1 \text{ to } 1.2 \text{ V}$$

The speed of the circuit is increased because it takes less time to charge up the stray capacitances of the load connected to the

output of the logic gate. As the output goes from high to low, the stray capacitances are charged by the collector current of $T2$ (with $T3$ and D nonconducting); when the output goes from low to high, this is done by the emitter current of $T3$ (with $T2$ nonconducting).

It is important that the improvement in the performance of a TTL gate with a totem-pole output driver is not accompanied by an increase in power dissipation under steady-state conditions, because at no-load a push-pull stage draws no power, irrespective of its state. A short-duration increase in consumption current can be observed at switching especially when the operating point traverses portion *III* of the transfer characteristic, within which both $T2$ and $T3$ are conducting simultaneously. To minimize the current amplitude at switching, a current-limiting resistor, $R3$, with a value of 200 to 300 Ω is placed in the collector lead of $T3$.

Current pulses existing in the supply leads might give rise to noise; to avoid it, it will be well-advised to place smoothing filters in the supply circuits. An optimal choice is to use a separate smoothing capacitor of 0.01 to 0.15 μF with a low stray inductance, for every two to eight logic gates.

AND-NOR Gates

In the logic gates examined so far, the second stage is a single-input NOT gate (inverter). If we increase the number of inputs on the second stage, we can expand the capabilities of logic gates and simplify, in some cases, the logic design of the associated digital system as a whole. If the first stage implements the AND function, the second stage must then implement the NOR function. The overall function instrument d by the resultant circuit may be written

$$y = \overline{\sum_{i=1}^{k} x_{i1} x_{i2} \cdots x_{iM_i}} \tag{5.38}$$

where k is the number of inputs to the second stage and M_i is the number of inputs of the first stage connected to the ith input of the second stage.

Logic gates mechanizing the switching function defined by Eq. (5.38) are known as AND-NOR gates. The respective logic symbol is shown in Fig. 5.35.

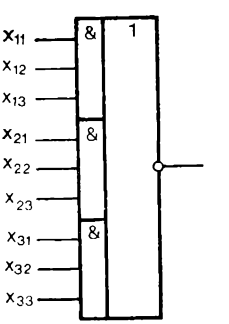


Fig. 5.35

In DTL gates, the number of inputs to the second stage can be increased by replacing the first coupling diode with a diode assembly and by connecting a diode AND gate to each input of the assembly (Fig. 5.36*a*, where $k = 2$ and $M_1 = M_2 = 2$). In TTL gates, an increase in the number of inputs to the second stage calls

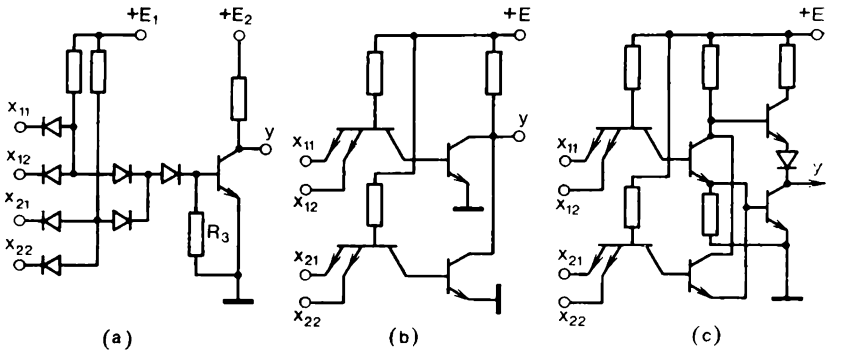


Fig. 5.36

for the use of parallel-connected inverting transistors (Fig. 5.36*b* and *c*).

5.5. CURRENT-MODE LOGIC GATES

These circuits are known as CML or ECL (emitter-coupled logic) gates. They have the highest speed of all other gates [24, 37], owing to the following factors.

(1) The transistors operate unsaturated, so they are free from the storage time;

(2) Owing to the inherent symmetry, there is practically no change in consumption current at change of state and no accompanying voltage surges in the supply circuits. Because of the low internal noise level, the difference between the high and low logic levels can be reduced, thereby cutting down the duration of transients associated with charging up the stray capacitances.

The circuit of a simplified single-input CML gate is shown in Fig. 5.37. Basically, it is a difference amplifier with a current generator I in the emitter circuit of the transistors. The base of $T2$ is connected to a d.c. voltage supply, E_B, and the base of $T1$ is driven by a control voltage, v_x. The state assumed by the transistors depends on the voltage between their bases, $V_x - E_B$. When $v_x - E_B$ is substantially positive, $T2$ is at cutoff, and all of the current flows through $T1$. When the sign of $v_x - E_B$ is reversed, $T1$ turns off, and all of the current flows through $T2$. At low values of $| v_x - E_B |$, both transistors are conducting. If the transistors have ideally identical characteristics, their collector currents become equal, $i_{C1} = i_{C2}$, at $v_x = E_B$.

Figure 5.38 shows plots of v_x as a function of v_{C1} and v_{C2}, and also a plot of v_{y2} (v_x) which controls the transfer characteristic of

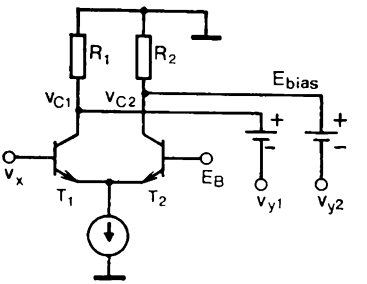


Fig. 5.37

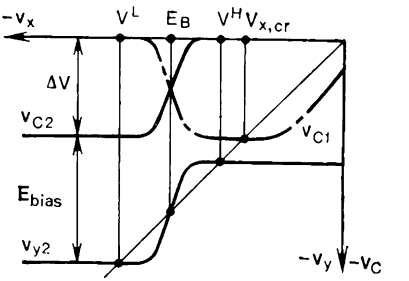


Fig. 5.38

the noninverting gate. The intersections of the v_{y2} (v_x) curve with the line of unit gain define the low and high logic levels, V^{L} and V^{H}.

Given small I_{CBO}, the collector voltage of a nonconducting transistor is zero very nearly. The maximum current, αI, of a conducting transistor determines the amplitude of collector voltage change at the time of switching:

$$\Delta V = \alpha I R_{\mathrm{C}} \approx I R_{\mathrm{C}} \tag{5.39}$$

In the active region, the collector potentials of $T1$ and $T2$ are higher than their base potentials. Because of this, the faithful reproduction of the input logic levels at the gate output calls for the use of an additional bias voltage source, E_{bias}, placed between the transistor collectors and the output terminals, as shown in Fig. 5.37. In practice, it may be assumed that at v_x corresponding to the low and high logic levels, the gate is fully switched (one of the transistors is driven to cutoff). We may also write that

$$V^{\mathrm{H}} = -E_{\mathrm{bias}}, \quad V^{\mathrm{L}} = -E_{\mathrm{bias}} - \Delta V$$

Consider the constraints that should be imposed on E_{bias} for the transistors to operate unsaturated. Because in the ON state of $T1$, its base voltage exceeds E_{B}, the risk of saturation exists above all for $T1$. For its collector junction to be reverse-biased, it is required that the voltage difference between the base and collector be negative. This condition for $T1$ is satisfied, if $v_{x,\max} \leqslant V_{x,\mathrm{cr}}$ (see Fig. 5.38). Noting that $v_{x,\max} = V^{\mathrm{H}} = -E_{\mathrm{bias}}$, and $V_{x,\mathrm{cr}} = v_{\mathrm{C1,min}} = -\Delta V$, the condition for the transistor to remain unsaturated may be written

$$E_{\mathrm{bias}} \geqslant \Delta V \tag{5.40}$$

Figure 5.39 shows the circuit of a CML (or ECL) gate adapted to implement a multi-input switching function. To this end, the input transistor has been replaced by a group of M parallel-connected transistors, and the bias voltage is supplied by the voltage drop

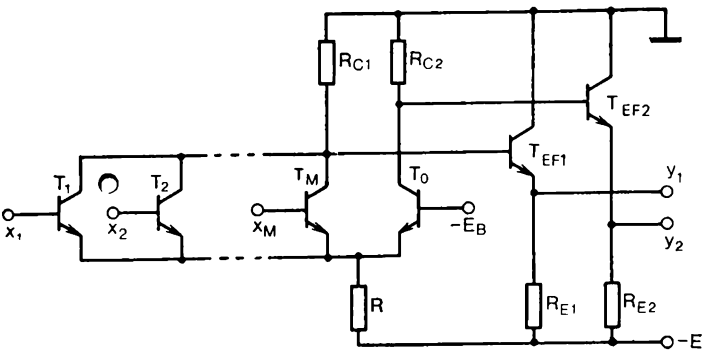


Fig. 5.39

across the emitter junctions of emitter followers $EF1$ and $EF2$ placed at the gate output. As a result, y_1 is a NOR function, and y_2 is an OR function.

Static noise immunity. Since the gate has two paraphase (complementary) outputs, it can be analyzed for static noise immunity, using the transfer characteristics of both an inverting and a noninverting type. However, the gain of a cascade of two CML gates is always greater than that of a single gate, and the noise margin referenced to the inverting output is larger than that referenced to the noninverting output. Accordingly, in calculating noise immunity, it is advisable to treat a CML gate as a noninverting one with a transfer characteristic v_{y2} (v_x).

To evaluate Max F, v_x must, in accordance with Table 5.2, be applied to all the inputs of the gate. In so doing, it is assumed that the input transistors are identical, that is, I_1 flowing in $R1$ is equally shared by $T1$, $T2$, . . . T_M and, as a consequence, v_x for a given I_1 is a minimum.

Let the emitter current of T_0 be denoted as i_{E2}, and the emitter current of the input transistor T_i $(i = 1, 2, \ldots, M)$ as i_{E1}. Then,

$$i_{E1} = (I - i_{EO})/M \tag{5.41}$$

The input voltage, v_x, differs from E_B by the difference in V_{BE} between the input transistor and T_0. Using Eq. (3.24) to evaluate the voltage across the emitter junction of a transistor, we get

$$v_x = E_B + m\varphi_T \,(\ln i_{E1} - \ln i_{E2}) \tag{5.42}$$

Solving Eqs. (5.41) and (5.42) simultaneously for i_{E2} gives

$$i_{E2} = I/\,(1 + z) \tag{5.43}$$

where

$$z = M \exp\,[(v_x - E_B)/m\varphi_T] \tag{5.44}$$

Using Eq. (5.43) and recalling that

$$v_{y2} = -(V_{BE,\,EF2} + 2\alpha i_{E2} R_{C2})$$

the upper bound on the transfer characteristic, Max F, can be written

$$v_{y2} = -[V_{BE,\,EF2} + \Delta V/(1+z)] \tag{5.45}$$

To evaluate Min F, we should, in accordance with Table 5.2, apply v_x to one of the inputs, leaving the other inputs low (in the logic 0 state). Now, switching the gate will cause only one input transistor to turn on. Therefore, the lower bound on the transfer characteristic is described by Eqs. (5.44) and (5.45) for $M = 1$. So, for Min F,

$$z = \exp\,[(v_x - E_B)/(m\varphi_T)] \tag{5.46}$$

To locate points A and B on the boundaries of the transfer characteristic, we evaluate the derivatives dv_{y2}/dv_x and equate them to zero to obtain

$$z^2 + 2(1 - 0.5\gamma)\,z + 1 = 0$$

where $\gamma = \Delta V/(m\varphi_T)$.

The minimum root of the equation corresponds to point A, and the maximum root to point B:

$$z_A = 0.5\gamma - 1 - \sqrt{0.25\gamma^2 - \gamma}$$
$$z_B = 0.5\gamma - 1 + \sqrt{0.25\gamma^2 - \gamma}$$

On substituting the value of z_A and z_B in Eqs. (5.44) through (5.46), we obtain the sought coordinates:

$$V_{y2A} = -\left[V_{BE,\,EF2} + \frac{2\Delta V}{\gamma(1-\sqrt{1-4/\gamma})}\right]$$
$$V_{xA} = E_B + m\varphi_T \ln(1/M)\,[\gamma/2 - 1 - (\gamma/2)\sqrt{1-4/\gamma}]$$
$$V_{y2B} = -\left[V_{BE,\,EF2} + \frac{2\Delta V}{\gamma(1+\sqrt{1-4/\gamma})}\right]$$
$$V_{xB} = E_B + m\varphi_T \ln[\gamma/2 - 1 + (\gamma/2)\sqrt{1-4/\gamma}]$$

Hence,

$$\Delta V_N^L = E_B + V_{BE,\,EF2} + \frac{2\Delta V}{\gamma(1-\sqrt{1-4/\gamma})} + m\varphi_T \ln[(1/M)(\gamma/2 - 1 - (\gamma/2)\sqrt{1-4/\gamma})] \tag{5.47a}$$

$$\Delta V_N^H = -E_B - V_{BE,\,EF2} - \frac{2\Delta V}{\gamma(1+\sqrt{1-4/\gamma})} - m\varphi_T \ln[\gamma/2 - 1 + (\gamma/2)\sqrt{1-4/\gamma}] \tag{5.47b}$$

The optimal value for E_B is chosen to obtain the equality $\Delta V_N^L = \Delta V_N^H$, whence

$$E_{B,\,opt} = -V_{BE,'EF2} - 0.5\Delta V + m\varphi_T \ln \sqrt{M} \qquad (5.48)$$

On substituting $E_B = E_{B,\,opt}$ in Eq. (5.47), we can evaluate the noise margin, $\Delta V_N = \Delta V_N^L = \Delta V_N^H$, of a CML gate at the optimal value of E_B:

$$\Delta V_N = \Delta V/2 - \frac{2\Delta V}{\gamma(1+\sqrt{1-4/\gamma})} - m\varphi_T \ln\left[\sqrt{M}\left(\gamma/2 - 1 + (\gamma/2)\sqrt{1-4/\gamma}\right)\right] \qquad (5.49)$$

As is seen from Fig. 5.40, a decrease in γ and an increase in M brings about a decrease in $\delta V_{N,opt} = \Delta V_N/\Delta V$. For each M, there is a minimum signal amplitude at which a logic gate still remains operable. At $M = 16$, the minimum amplitude is $10\ m\varphi_T = 260$ to 350 mV.

The above constraint on the minimum voltage difference between the logic levels is of fundamental nature and holds for all logic gates based on bipolar transistors.

For a CML gate, the signal amplitude ΔV, as follows from Eq. (5.39), does not exceed 0.6 or 0.7 V. The value of γ varies with m and usually ranges between 16 and 28. As is seen from Fig. 5.40, for typical values of ΔV at $M = 16$, the ratio of noise margins remains greater than 0.2.

Power dissipation. The total power dissipation P is the sum of the powers drawn by the current switch, P_{sw}, and by the emitter followers, P_{EF}. Recalling Eq. (5.39), we get

$$P_{sw} = \Delta V E/R_C \qquad (5.50)$$

and

$$P_{EF} = E(I_{EF1} + I_{EF2}) = \frac{2E^2 + E(V^H + V^L)}{R_E} = (2E^2/R_E)\left(1 - \frac{2e_{EO} + \Delta V}{2E}\right) \qquad (5.51)$$

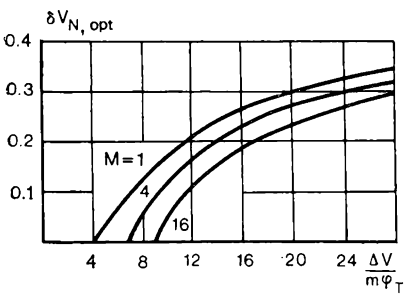


Fig. 5.40

Dynamic characteristics. Because of the low impedance presented by the external circuits to the transistors of a CML (ECL) gate, the control (driving) currents and, as a consequence, transient response are substantially affected by the nonlinearity of the input impedances of the transistors. This handicaps a precise calculation of dynamic characteristics for CML (ECL) gates and calls for resort to computers.

An approximate estimate for the propagation delay time can be derived on making the following assumptions:

(1) the sluggish response of the gate is mainly due to re-charging the barrier and stray circuit capacitances, whereas the time required to switch

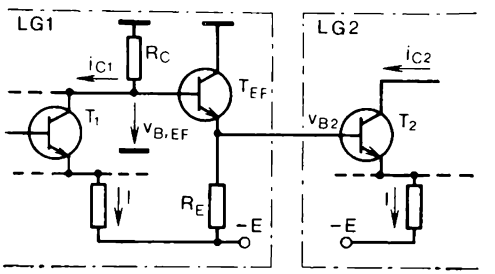


Fig. 5.41

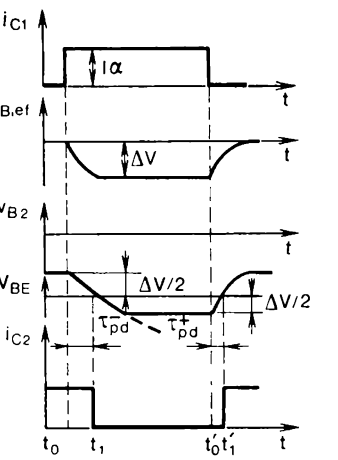


Fig. 5.42

the emitter currents in the transistors may be neglected;

(2) the transfer function of a CML gate is close to a step function with a threshold at half the peak value of the signal.

The above assumptions are correct for high-frequency drift (integrated-circuit) transistors and for a signal amplitude of at least 0.4 V.

To define the dynamic characteristics of a CML gate, it will suffice to trace the passage of the signal from *LG*1 to *LG*2 (from the collector of *T*1 to the collector of *T*2 in Fig. 5.41), with transistors *T*1 and *T*2 forming the arms of the current switches in *LG*1 and *LG*2, respectively.

The input signal at time t_0 is a positive step, αI, of the current i_{C1}, and at time t_0', a negative step, $-\alpha I$, of the same current. The step change in the collector current of *T*1 produced by charging up the stray capacitances leads to a gradual change in the base voltages, $v_{B,EF}$ and v_{B2}, of the transistors T_{EF} and *T*2 (Fig. 5.42) and to a step change in i_{C2} at the instant when the absolute change in v_{B2} becomes equal to 0.5 ΔV. The time intervals from t_0 to t_1 and from t_0' to t_1' define the respective delay times, τ_{pd}^- and τ_{pd}^+.

The positive step change in i_{C1} results in a rapid decrease in $v_{B,EF}$, so that the emitter follower is turned off, and the total equivalent load capacitance C_Σ is recharged via R_E, as shown in Fig. 5.43. C_Σ is the sum of the C_e and C_c of the input transistors in the component loading *LG*1, and the circuit (wiring) capacitance C_w. Owing to the high value of R_E, it is legitimate to ignore the base bulk resistance of the input transistors.

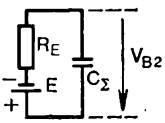


Fig. 5.43

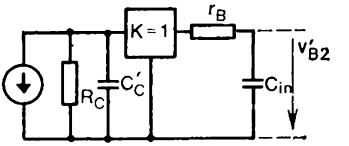


Fig. 5.44

Referring to the equivalent circuit of Fig. 5.43, the value of v_{B2} varies exponentially

$$v_{B2}(t) = (E - e_{EO}) \exp\left(-\frac{t-t_0}{R_E C_\Sigma}\right) - E$$

Equating $v_{B2}(t)$ to the threshold voltage ($V^H - \Delta V/2$), we get

$$\tau_{pd}^- = -R_E C_\Sigma \ln\left(1 - \frac{1}{2}\frac{\Delta V}{E - e_{EO}}\right) \approx \frac{R_E C_\Sigma \Delta v}{2(E - e_{EO})} \qquad (5.52)$$

When the input is a negative step change in i_{C1}, the current of the emitter follower rises.

With an accuracy sufficient for approximate calculations, we may deem that an emitter follower has a voltage transfer ratio (voltage gain) of unity, a zero output resistance, and an infinite input resistance. Owing to the low value of output resistance, we have to take into account the increased effect of the base bulk resistance r'_B of the input transistors on charging, but, at the same time, we may neglect the effect of the additional load.

In the path traversed by a negative step change in i_{C1} (Fig. 5.44), the total capacitance C'_C between the output of the current switch and the common return is the sum of the wiring capacitance C_ω and the collector transition (barrier) capacitance C_c. If the ECL (CML) gate is an integrated circuit, to this we should add C_s and $C_{R,C}$ that appear between the circuit components and the substrate:

$$C'_C = M(C_c + C_s) + C_e + C_{R,C} + C_w$$

The input capacitance C_{in} is the sum of the C_c and C_e of the input transistor.

Referring to the circuit of Fig. 5.44 and recalling Eq. (5.39), we obtain

$$v'_{B2\sim}(s) = \frac{\Delta V}{s(1 + sR_C C'_C)(1 + sr'_B C_{in})} \qquad (5.53)$$

In practice, it is usual that $R_C C'_C \gg r'_B C_{in}$. Therefore, it is legitimate to replace $(1 + sr'_B C_{in})$ by an exponential term, $\exp(sr'_B C_{in})$. Then the original time function corresponding to Eq. (5.53) may

be written

$$v'_{B2\sim}(t) = \Delta v\left[1 - \exp\left(\frac{t - r'_B C'_C - t'_0}{R_C C_c}\right)\right]$$

Solving $v_{B_2}(t'_1) = 0.5\Delta V$ gives

$$\tau^+_{pd} = r'_B C_{in} + R_C C'_C \ln 2 \approx r'_B C_{in} + 0.7 R_C C'_C \qquad (5.54)$$

The propagation delays τ^-_{pd} and τ^+_{pd} are related to the power that a CML gate draws from its power supply:

$$\tau^-_{pd} = \frac{E\Delta V C_\Sigma\left(1 - \dfrac{2e_{EO} + \Delta V}{2E}\right)}{P_{EF}(1 - e_{EO}/E)} \approx \frac{E\Delta V C_\Sigma}{P_{EF}}(1 - \Delta V/2E) \qquad (5.55)$$

$$\tau^+_{pd} = r_B C_{in} + \frac{0.7 E\Delta V C'_C}{P_{sw}}$$

The average delay, $\bar{\tau}_{pd} = (\tau^-_{pd} + \tau^+_{pd})/2$, is given by

$$\bar{\tau}_{pd} = 0.5 r_B C_{in} + (E\Delta V/2P)\left[0.7 C'_C/\gamma_P + \frac{C_\Sigma}{1 - \gamma_P}(1 - \Delta V/2E)\right] \qquad (5.56)$$

where $\gamma_P = P_{sw}/P$.

It follows from Eq. (5.56) that a higher speed of the logic gate leads to an increased power dissipation. At a constant P, the average delay decreases with a decrease in the supply voltage E, the signal amplitude ΔV, and C'_C and C_Σ. There is an optimal value of γ_P such that for given C'_C, C_Σ, E, and ΔV, the average delay is a minimum:

$$\gamma_{P,\,opt} = 1 + \sqrt{1.4(1 - \Delta V/2E)(C_\Sigma/C'_C)}$$

The average delay for state-of-the-art CML (ECL) gates ranges from 0.5 to 5 ns at a power dissipation of 35 to 70 mW per stage.

5.6. MOSFET LOGIC GATES

MOSFET (or MOS-transistor) logic gates account for a large proportion of the components used in LSI digital networks. This is because they are simple to manufacture and small in size, and draw little power, so that a large number of devices can be fabricated on a single substrate.

The name MOSFET or MOS-transistor applies to an arrangement in which a metal gate electrode is separated by an oxide layer from the semiconductor channel. This metal-oxide-semiconductor (MOS) structure allows an electric field to affect the channel, if an external potential is applied between the gate and the substrate.

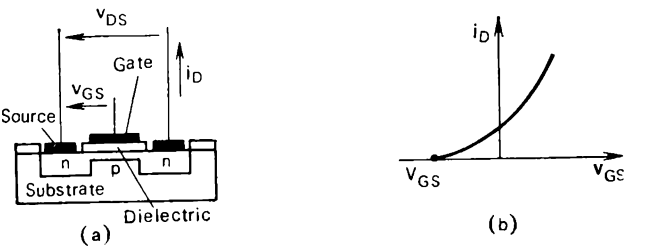


Fig. 5.45

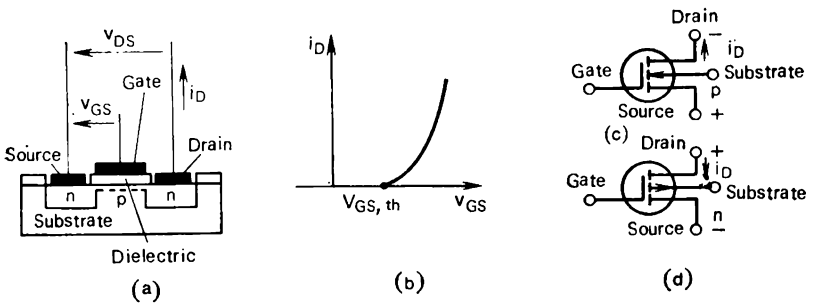


Fig. 5.46

There are two basic types of MOS transistor: the depletion type and the enhancement type. The structure of a depletion-mode MOS-transistor is shown in Fig. 5.45*a*. The channel has the same type of conduction as the drain and source regions and opposite to that of the substrate (body). In such a transistor, the drain current can flow at zero voltage between the gate and source, v_{GS} (Fig. 5.45*b*).

The structure of an enhancement-mode MOS-transistor is shown in Fig. 5.46*a*; it differs from the depletion device in that no channel exists at $v_{GS} = 0$. Therefore, the drain-source circuit consists in effect of two diodes connected in opposition, so that the drain current is zero. The channel is induced under the action of the voltage applied to the gate. The minimum voltage required to induce the channel is called the *threshold voltage*, $V_{GS,\,th}$ (Fig. 5.46*b*). The cut-in voltage has the same polarity as the drain supply voltage, being positive for *n*-channel MOS-transistors (Fig. 5.46*c*) and negative for *p*-channel MOS-transistors (Fig. 5.46*d*). Because an enhancement-mode MOS-transistor can be turned off, with v_{DS} and v_{GS} being of the same polarity, the stages can be direct-coupled, and MOS-transistors can conveniently be used to build common-drain logic circuits.

The *V-I* characteristics of a MOS-transistor, obtained on making some simplifying assumptions [20], may be written analytically as

$$i_D = \begin{cases} \nu\,[2v_{DS}(v_{GS} - V_{GS,\,th}) - v_{DS}^2] \\ \qquad \text{for } v_{DS} < v_{GS} - V_{GS,\,th} \qquad (5.57a) \\ \nu\,(v_{GS} - V_{GS,\,th})^2 \\ \qquad \text{for } v_{DS} \geqslant v_{GS} - V_{GS,\,th} \qquad (5.57b) \end{cases}$$

where ν is a characteristic parameter depending on the transistor geometry and the electrophysical properties of the semiconductor material; it has the dimensions of $\mu A/V^2$.

Within the region described by Eq. (5.57a), the drain current i_D markedly depends on v_{DS}, because of which this portion is referred to as the ohmic (also called the triode or nonsaturation) region. The region described by Eq. (5.57b), i_D is theoretically independent of v_{GS}, and it is called the saturation (pentode) region.

MOS-transistor inverter. Figure 5.47 shows the circuits of two basic types of MOS-transistor inverters, one with a saturated load and the other with an unsaturated load.

An MOS-transistor inverter with a saturated load (Fig. 5.47*a*) consists of a driver transistor $T1$ and a load transistor $T2$. Both have the same type of channel. An insight into the operation of the inverter can be gleaned from the *V-I* characteristics: $i_D(v_y)$ for the driver transistor and $i_L(v_y)$ for the load transistor, plotted on the basis of Eqs. (5.57). In the simplest case, the gate of $T2$ can be connected to the supply E (the dashed line in Fig. 5.47*a*). Then, for the load transistor $v_{GS} = v_{DS} = E - v_y$, and it operates within the saturation region described by

$$i_D = i_L = \nu_L (E - V_{GS,\,th}\, v_y)^2 \qquad (5.58)$$

corresponding to load curve *1* in Fig. 5.48*a*. The performance of the inverter can be improved by connecting the gate of $T2$ to a separate

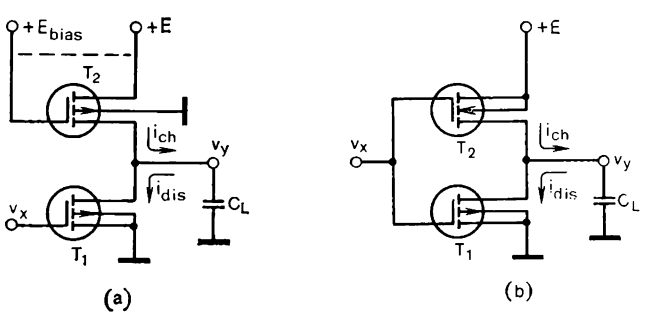


Fig. 5.47

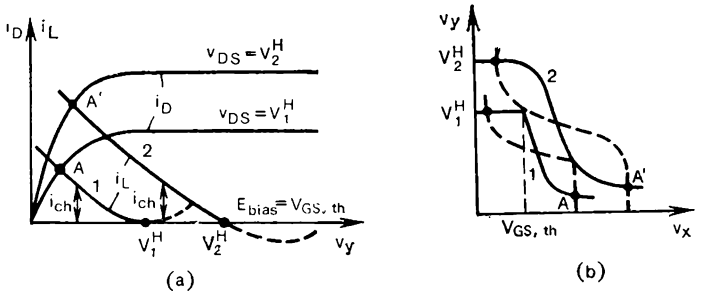


Fig. 5.48

bias supply E_{bias}, so that with $E_{bias} > E + V_{GS,\,th}$ the load transistor operates within the saturation region. In accord with Eq. (5.57b), the V-I characteristic can be written

$$i_L = \nu_L (E - v_y)(2E_{bias} - 2V_{GS,th} - E - v_y) \qquad (5.59)$$

corresponding to load curve *2* in Fig. 5.48*a*.

The transfer characteristics of a saturated-load MOSFET inverter are shown in Fig. 5.48*b*. So long as $v_x < V_{GS,th}$, the driver transistor is off. In a steady state, $i_L = 0$. Therefore, in accord with Eqs. (5.58) and (5.59), the output voltage of an unbiased MOSFET inverter will be $E - V_{GS,th}$, whereas in a biased MOSFET inverter it will be $v_y = E$. These voltages fix the high voltage level V^H corresponding to a logic 1 in positive logic. The low voltage level V^L is determined by the output voltage at point A or A' on the transfer characteristics.

The speed of a MOSFET inverter depends on the charging and discharging of the load capacitance C_L which is the sum of the wiring (circuit) capacitance and the input capacitance of the succeeding logic gates. C_L is charged up via the load transistor. To obtain small values for V^L in saturated-load MOSFET inverters, $T2$ should be chosen to have a high-resistance channel (a low value of ν_L). This tends to slow down the charging of C_L.

A biased MOSFET inverter has a more linear load characteristic, and the current charging the load capacitance decreases as v_y approaches V^H, at a slower rate than in an unbiased inverter. As a result, a biased MOSFET inverter has a higher operating speed.

A far better performance is shown by an unsaturated-load MOSFET inverter (Fig. 5.47*b*). Its load is formed by an *n*-channel and a *p*-channel transistor. When the inverter input is high, $T2$ turns off and $T1$ turns on, thereby speeding up the discharge of C_L to zero. Similarly, when the inverter input goes low, C_L rapidly charges up via the ON transistor $T2$, whereas $T1$ remains OFF (Fig. 5.49).

By connecting driver transistors in parallel, series, or series-parallel, MOS-transistors can be arranged into multi-input logic

gates. In structure, a nonlinear-load multi-input logic gate using parallel-connected MOS-transistors (Fig. 5.50*a*) is similar to the common-emitter bipolar-transistor logic gate examined in Sec. 5.3. With positive logic and *n*-channel MOS-transistors, such a circuit implements the NOR function. To implement the NAND function, the MOS-transistors are connected in series (Fig. 5.50*b*). It should be pointed out, however, that the lower arm of the network dividing the voltage generated at $y = 0$ contains a series chain of M transistors, and the low-level voltage V^{L} in the NAND gate is higher than in the NOR gate. To bring down V^{L} in a NAND gate, one has to use low-resistance transistors occupying a sizeable area on the substrate. The more complex logic functions are mechanized by arranging MOS-transistors in parallel, series, and series-parallel.

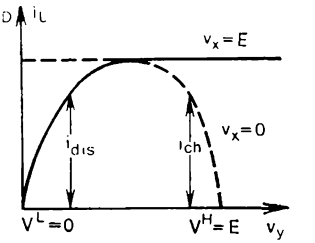


Fig. 5.49

When *p*-channel and *n*-channel MOS-transistors are combined, the resulting circuitry is termed a complementary MOS, or CMOS. In a CMOS logic gate, the driver transistors should be connected in series-parallel in both the upper and lower arms of the divider. The resulting logic gate may be depicted as a series connection of two switches (Fig. 5.51*a*). Let the state of the upper switch be denoted by z_u and that of the lower switch by z_l. Also assume that when a switch is closed, it is in a logic 1 state and that, when open, it is in a logic 0 state. At $y = 1$, the upper switch is closed and the lower switch is open, whereas at $y = 0$ the upper switch is open and the lower switch is closed. Therefore, in order to realize the function $y = F(x)$, we must require that the values of z_u be the same as of F and the values of z_l the same as those of $\bar{F}$. Closure of the switch in the lower arm

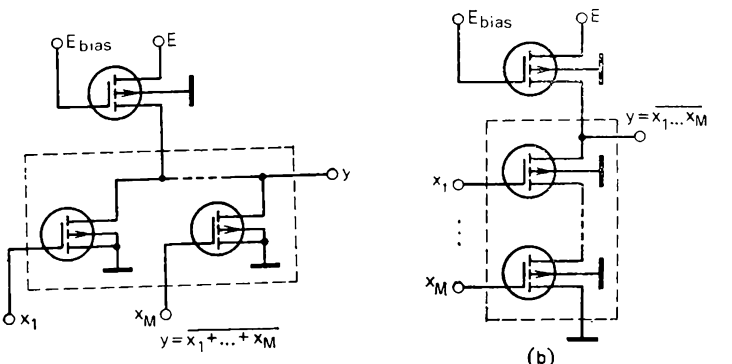


Fig. 5.50

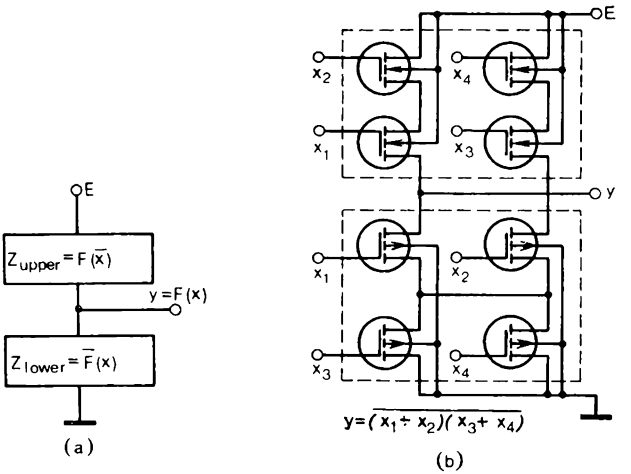


Fig. 5.51

occurs when a high-level voltage is applied to its gate, whereas closure of the switch in the upper switch takes place when a low-level voltage is applied to its gate. Because of this, it is convenient to take the inverted input variables as the arguments of the function z_u, and the input variables themselves as the arguments of the function z_l. On writing Boolean expressions for z_u and z_l, we can see that each AND operation can be implemented by a series connection of MOS-transistors, and each OR operation by a parallel connection. As an example, Fig. 5.51*b* illustrates the implementation of the function

$$y = \overline{(x_1 + x_2)\,(x_3 + x_4)}$$

for which

$$z_l = (x_1 + x_2)\,(x_3 + x_4)$$

and

$$z_u = \overline{(\bar{x}_1 + \bar{x}_2)\,(\bar{x}_3 + \bar{x}_4)} = x_1x_2 + x_3x_4$$

REVIEW QUESTIONS AND PROBLEMS

5.1. Draw up a truth table for the switching functions $y_1 = x_1 + x_2$, $y_2 = x_1x_2$, $y_3 = x_1 \oplus x_2$.

5.2. Construct logic networks using AND, OR and NOT gates to implement the switching functions $y_1 = x_1\bar{x}_2 + \bar{x}_1x_2$, $y_2 = \overline{x_1 + x_2} + x_1\bar{x}_3$, $y_3 = x_1\,(x_2 + \bar{x}_3\bar{x}_4) + x_2x_4$.

5.3. Draw up a truth table for the switching functions $y_1 = \overline{x_1 + x_2}$, $y_2 = \overline{x_1 x_2}$.

5.4. Construct logic networks using NAND and NOR gates to implement the switching functions in Problem 5.2.

5.5. List methods for representing binary variables by electric signals.

5.6. Plot the transfer characteristic for a noninverting logic gate. Using the transfer curve thus plotted, determine the logic 0 and 1 levels for positive and negative logic.

5.7. Determine logic 0 and 1 levels for a noninverting logic gate, considering the likely spread in its transfer characteristic.

5.8. Plot the transfer characteristic for an inverting logic gate. Using the curve thus plotted, determine the logic 0 and 1 levels. List the conditions to be met by the transfer characteristic of an operable logic gate.

5.9. How can the noise margin be found from the static transfer characteristic of a logic gate? Point out the difference in working this problem for an inverting and a noninverting logic gate.

5.10. Referring to the condition for optimal static noise immunity, calculate the values of the resistors in the logic gate of Fig. 5.18 for which $\overline{P} = 20$ mW, $M = 4$, $N = 6$, and $E_C = 10$ V. The parameters of a simplified piecewise-linear transistor model are: $e_{EO} = 0.7$ V, $V_{CE\,(sat)} = 0.1$ V, $I_{CBO,\,max} = 1$ μA, $\beta_{min} = 60$.

5.11. Work Problem 5.10 for the RTL gate of Fig. 5.25, for which $\overline{P} = 2$ mW, $N = M = 8$, and $E_C = 4$ V.

5.12. Work Problem 5.11, using an exponential transistor model for which the parameters are: $\beta_{min} = 60$, $\beta_I = 1$, $m = 1.2$, $r'_B = 100\ \Omega$. Assume that for $V_{CE} = 4$ V, $I_{CBO,\,max} = 1$ μA and that for $I_{B(sat)} = I_{C(sat)} = 5$ mA, $V_{BE(sat)} = 1.1$ V, and $V_{CE(sat)} = 0.1$ V.

5.13. For the DTL gate of Fig. 5.29, $E_a = 6$ V, $E_C = 3$ V, $E_{bias} = 0$, $R_B = R_a = 20$ kΩ, $R_C = 10$ kΩ. The transistor parameters are: $e_{EO} = 0.6$ V, $\beta = 60$. Calculate I_a, I_{bias}, and $I_{R,\,C}$, for the two states of the logic gate and determine the saturation coefficient of the transistor for $N = 5$.

5.14. For the DTL gate of Fig. 5.29, $E_a = 6$ V, $E_C = 3$ V, $E_{bias} = 0$, $R_{a,\,max} = 30$ kΩ, $R_{a,\,min} = 20$ kΩ, $R_{C,\,max} = 15$ kΩ, $R_{C,\,min} = 10$ kΩ. The transistor parameters are: $\beta_{min} = 50$, $e_{EO,\,max} = 0.8$ V. Calculate the value of R_B at which the logic gate has a maximum noise immunity.

5.15. List the advantages of TTL gates over DTL gates.

5.16. List the factors governing the noise margin of TTL gates using a simple inverter. Compare with DTL gates.

5.17. Plot the transfer characteristic for a TTL gate with a totem-pole output driver. Determine the noise margin of the gate from the resultant curve graphically.

5.18. List the factors ensuring high speed for CML (ECL) gates.

5.19. Draw the circuit of a CML (ECL) gate and explain the function of the emitter followers.

5.20. List ways and means of optimizing the static and dynamic characteristics of a CML (ECL) gate.

6 BISTABLE CIRCUITS

6.1. GENERAL

A bistable circuit, as its name implies, is a switching circuit with two stable states. It can be triggered from one state to the other by applying an input voltage via a suitable trigger circuit. Most frequently, bistable circuits (often called simply "bistables") are of the static (d.c. or level) type in which the stable states differ in output voltage level.

Static bistable circuits generally serve two distinct functions, namely: as pulse shapers and in logic networks. In the former case, they are more usually called *bistable multivibrators*, and in the latter, *flip-flops*. Bistable multivibrators produce rectangular pulses or even square waves with a standard amplitude and short rise and fall times practically independent of the rate of change of the input signal. Flip-flops serve as binary memory (storage) cells, one-bit delay circuits, and scale-of-2 counters or binaries (counters which produce one output pulse for every two input pulses).

Most frequently, a static bistable circuit can be obtained by applying positive feedback to a d.c. amplifier. The manner in which stable states are produced in this case is illustrated in Fig. 6.1. The plot of $v_{out} = f(v_{in})$ is the transfer characteristic of the loaded d.c. amplifier, and the $v_{out} = v_{in}$ line is the feedback characteristic (it is assumed that the output is direct-coupled to the input). The intersections of the transfer curve with the feedback line define the conditions for the circuit to be at equilibrium. Two stable states will be obtained, if the feedback line cuts the transfer curve at three points, as shown in Fig. 6.1. Points *A* and *B* where the gain (with the feedback loop open) is less than unity correspond to stable states, whereas point *C* corresponds to a temporary state, because the gain at that point exceeds unity. Even a minute deviation from point *C* (because of, say, noise in the circuit) will be amplified over the feedback loop and cause the bistable circuit to

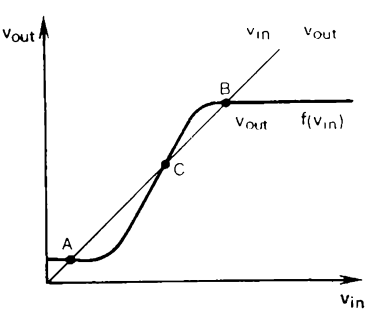


Fig. 6.1

go from any one of the two stable states at a progressively increasing rate.

Because of this, the bistable circuit may be treated as regenerative. The rate of regeneration is governed by the time lag (inertia) associated with the stray capacitances existing in, and the semiconductor devices used by, the d.c. amplifier.

As often as not, the d.c. amplifier in a bistable circuit is made with logic gates. From a comparison of Figs. 6.1 and 5.5 it is seen that the condition for the existence of two stable states is analogous to that for the reproduction of logic levels by a logic gate. In other words, when appropriately chosen and interconnected, logic gates will make an operable bistable circuit.

One of the principal performance characteristics of bistable circuits is the *maximum switching frequency*, F_{max}. It is defined as the maximum number of valid changes of state per second and is measured in hertz, Hz. The minimum time interval, T_{min}, between adjacent trigger pulses that cause a bistable circuit to change state is named its *resolving* or *recovery time*. Obviously, $T_{min} = 1/F_{max}$.

Among the other characteristics related to the performance of a bistable circuit are the shape and amplitude of output pulses, trigger-circuit parameters, maximum load, power dissipation, and size.

Apart from static bistable circuits, there are dynamic ones, usually operated as clocked circuits. In them, a binary signal can be delayed for one bit time by sorting it in reactive (L and C) components. Dynamic bistable circuits can readily be arranged into digital networks of any arbitrary configuration. In IC form, dynamic bistable circuits can most readily be made with MOS structures in which the reactive storage element is the stray capacitance between the gate and drain (see Chap. 13 for more detail).

Two stable states can be obtained by placing a nonlinear one-port with an S- or N-shaped V-I characteristic in a resistive circuit. For this arrangement to have two stable states, it is important that the load line should cut the V/I curve at three points. The intersection with the downward portion of the V/I curve corresponds to a temporary state and is due to the internal feedback existing in the device itself (see Chap. 12).

6.2. UNSYMMETRICAL BISTABLE MULTIVIBRATORS

An unsymmetrical bistable multivibrator has one input and a transfer characteristic of the form known in Fig. 6.2*a*. Portion *2* of the characteristic corresponds to bistable operation. To make our discussion more specific, let the state in which v is equal to the high logic level, V^H, be the 1 state, and the state in which $v = V^L$ be the 0 state. Only one logic state can exist within portion *1* or *3*. The arrows show that the circuit goes from the 0 to the 1 state at the

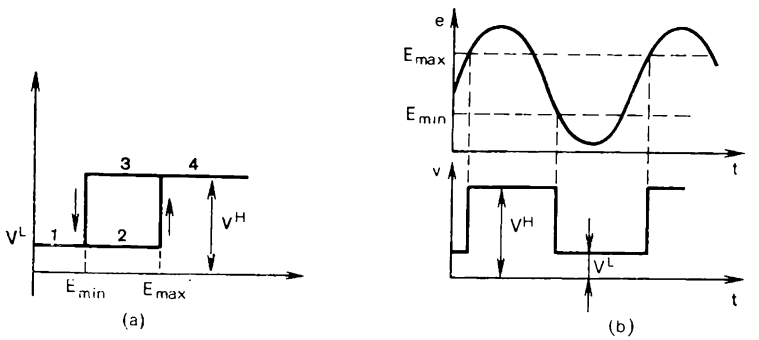


Fig. 6.2

instant when the positive-going input voltage e passes through a boundary point labelled E_{max}, and from the 1 to the 0 state when the negative-going voltage e passes through another boundary point labelled E_{min}. The values of E_{max} and E_{min} respectively define the *set* and *reset* threshold.

An unsymmetrical bistable multivibrator usually operates as a threshold circuit which responds to a particular signal level (or a particular pulse amplitude) or as a rectangular-pulse generator which derives rectangular pulses of a specified amplitude from a sinewave or any other nonrectangular input waveform (Fig. 6.2*b*).

Basically, unsymmetrical bistable multivibrators come in two modifications, one using emitter-coupled transistors and the other a differentiating opamp.

Emitter-Coupled Bistable Multivibrators

In an emitter-coupled bistable multivibrator (Fig. 6.3), the voltage from the collector load of $T1$ in the d.c. amplifier is taken to the base of $T2$ via a resistor voltage divider R_C/R_B, the upper arm of which is shunted by a speed-up (commutating or memory) capacitor C. The feedback loop is completed by tieing together the emitters of $T1$ and $T2$. The drive voltage e is applied to the base of $T1$, and the output voltage v is taken from the collector of $T2$.

Consider the operation of the bistable multivibrator, assuming that the drive voltage e is slowly varying from zero to E_C. Suppose also that initially $e = 0$, so that $T1$ is OFF and $T2$ is in the active region*. The cutoff voltage on the emitter of $T1$ is due to the emitter current

* Alternatively, $T2$ may be at saturation when $T1$ is OFF [4]. The set and reset levels will then be affected by the load conductively coupled to the $T2$ collector.

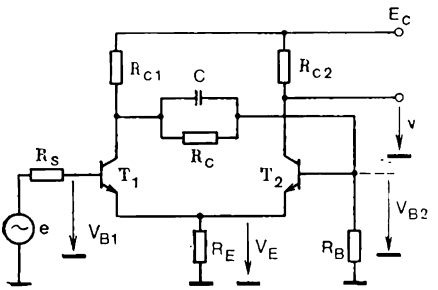


Fig. 6.3

of $T2$ flowing through R_E. The collector voltage of $T2$ defines the low logic level, V^L, of output voltage.

As e rises, an instant is reached when $e(t) = E_{max}$, and $T1$ turns ON, so that the current flowing through R_E is shared between $T1$ and $T2$. Since both transistors are in the active region, the feedback loop is completed. Now the feedback loop gain exceeds unity (which is the condition for bistable operation), so that the current through $T1$ builds up and the current through $T2$ collapses at a progressively increasing rate. Finally, $T1$ turns OFF, and the output voltage goes abruptly from V^L to $V^H = E - I_{CBO}R_{C2}$.

In the 1 state, $T1$ can be either at saturation or in the active region, depending on the value of e. In the former case, the collector voltage of $T1$ is equal to its emitter voltage, so the condition for $T2$ to turn OFF is satisfied always, owing to the presence of the voltage divider R_c/R_B.

A fall in e first moves $T1$ into the active region, then (at $e(t) = E_{min}$) drives $T2$ into conduction. As a result, the feedback loop is completed and a regenerative process is initiated, which terminates in turning off $T1$ and resetting the circuit (bringing it back to the zero state).

Operability condition. For an emitter-coupled bistable multivibrator to have two stable states, it is essential that E_{max} and E_{min} at which $T1$ turns on as e increases and $T2$ turns on as e decreases be such that

$$E_{max} > E_{min} \tag{6.1}$$

As will be shown later, this requirement is equivalent to having a feedback loop gain of more than unity, so this is a sufficient condition for bistable operation (see Fig. 6.3). In addition to satisfying Eq. (6.1), the normal operation of the emitter-coupled bistable multivibrator requires $T2$ to be unsaturated. A sufficient condition for this requirement to be met is

$$v_{C2} = V^L \geqslant v_{B2} = v_E + e_{EO2} \tag{6.2}$$

When $T1$ is OFF (the 0 state), the emitter-coupled bistable multivibrator can be depicted by the equivalent circuit of Fig. 6.4*a*. On finding v_{B1} and v_E from the equivalent circuit and assuming that $T1$

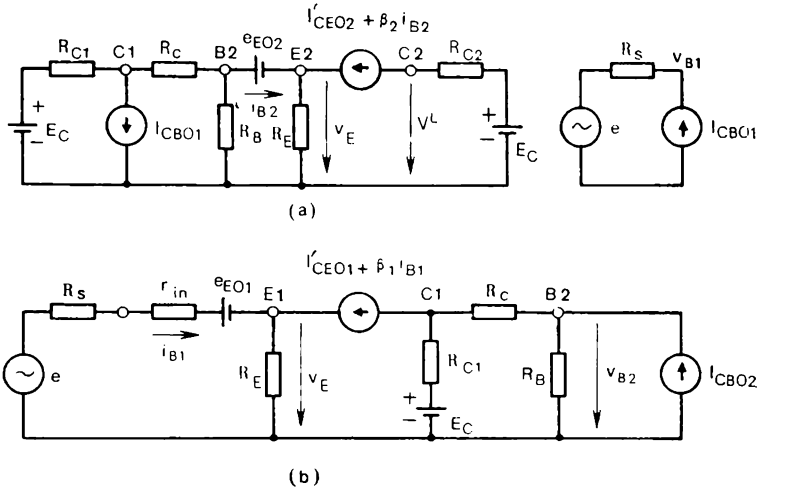

$I'_{CEO2} + \beta_2 i_{B2}$
R_{C1} C1 R_c B2 e_{EO2} E2 C2 R_{C2}
R_s v_{B1}
E_C I_{CBO1} i_{B2} R_B R_E v_E V^L E_C
e I_{CBO1}
(a)
$I'_{CEO1} + \beta_1 i_{B1}$
R_s r_{in} e_{EO1} E1 C1 R_c B2
e i_{B1} R_E v_E R_{C1} E_C R_B v_{B2} I_{CBO2}
(b)

Fig. 6.4

turns ON at $v_{B1} = v_E + e_{EO1}$, the set level is found to be

$$E_{max} = \frac{E_C - I_{CBO1}R_{C1} + I_{CBO2}(R_{C1}+R_{C2})}{1+\frac{R_{C1}+R_c}{R_B}\left(1+\frac{R_B}{R_E(1+\beta_2)}\right)} - \frac{e_{EO2}}{1+\frac{R_B(R_{C1}+R_c)}{R_E(1+\beta_2)(R_{C1}+R_c+R_B)}} + e_{EO1} - I_{CBO1}R_s \qquad (6.3a)$$

As is seen, to obtain a stable set level, it is essential that $I_{CBO1}(R_{C1} + R_c) \ll E_C$, $I_{CBO1}R_s \ll E_{max}$, and $R_B/R_E(1 + \beta_2) \ll 1$. The last inequality minimizes the effect of β_2, on the one hand, and permits e_{EO1} and e_{EO2} to cancel each other, on the other. Therefore, Eq. (6.3a) may be re-cast as

$$E_{max} \approx E_C\left[1+\frac{R_{C1}+R_c}{R_B}\left(1+\frac{R_B}{R_E(1+\beta_2)}\right)\right]^{-1} \qquad (6.3b)$$

Using the equivalent circuit of Fig. 6.4*a*, the condition for $T2$ to be unsaturated may be written as

$$R_{C1} < R_E(E_C/E_{max} - 1)$$

When $T2$ is OFF (the 1 state), the equivalent circuit of the emitter-coupled bistable multivibrator takes the form shown in Fig. 6.4*b*. On finding v_E and v_{B2} from the equivalent circuit and assuming that

$T2$ turns ON at $v_{B2} = v_E + e_{EO_2}$, the reset level is found to be

$$E_{min} = \frac{E_C + I_{CBO1}R_{C1} + I_{CBO2}(R_{C1} + R_c) - e_{EO2}[1 + (R_{C1} + R_c)/R_B]}{1 + (R_{C1} + R_c)/R_B + \alpha_1 R_{C1}/R_E}$$

$$\times \left[1 + \frac{R_s + r_{in}}{R_E(1 + \beta_1)}\right] + e_{EO1} - I_{CBO1}(R_s + r_{in}) \qquad (6.4a)$$

Because the reverse (leakage or cutoff) collector current of $T1$ produces a stronger effect ($I_{CEO} \gg I_{CBO}$) and also because e_{EO_1} and e_{EO_2} do not cancel each other completely, the stability of E_{min} is, on the whole, lower than that of E_{max}. In calculations not related to stability, Eq. (6.4a) may be simplified by decreasing the infinitesimals:

$$E_{min} \approx \frac{E_C[1 + R_s/R_E(1 + \beta_1)]}{1 + (R_{C1} - R_c)/(R_B + \alpha_1 R_{C1}/R_E)}$$

On substituting Eqs. (6.3b) and (6.4b) in Eq. (6.1), the condition for the normal operation of an emitter-coupled bistable multivibrator can be written as

$$\alpha_1(1 + \beta_2) R'_{C1}/(R_{C1} + R_c)\left[1 + \frac{R_s}{R_E(1 + \beta_1)}\left(1 + \frac{R_E(1 + \beta_2)}{R_B}\right)\right] > 1 \qquad (6.5)$$

The left-hand side of Eq. (6.5) defines the feedback loop gain, so Eq. (6.5) and, as a consequence, Eq. (6.1) is a sufficient condition for bistable operation.

Unsymmetrical Bistable Multivibrator Using a Differential Opamp

As already noted in Chap. 4, a differential operational amplifier has a nearly relay (ON-OFF) response. To obtain a hysteresis response, a feedback loop is connected to the noninverting input of the opamp (Fig. 6.5*a*). In the active region, the feedback loop gain K_p is very high:

$$K_p = \gamma K_0 \gg 1 \qquad (6.6)$$

where $\gamma = R_2/(R_1 + R_2)$.

Therefore, in locating points A and B which define stable states, the differential opamp may be deemed to have an ON-OFF transfer characteristic, with the vertical portion passing through zero (v_2 at the noninverting input is zero) when the voltage at the inverting input is zero, $e = 0$ (Fig. 6.5*b*). Referring to Fig. 6.5*b*, the ordinates of points A and B are the intersections of the horizontal portion of

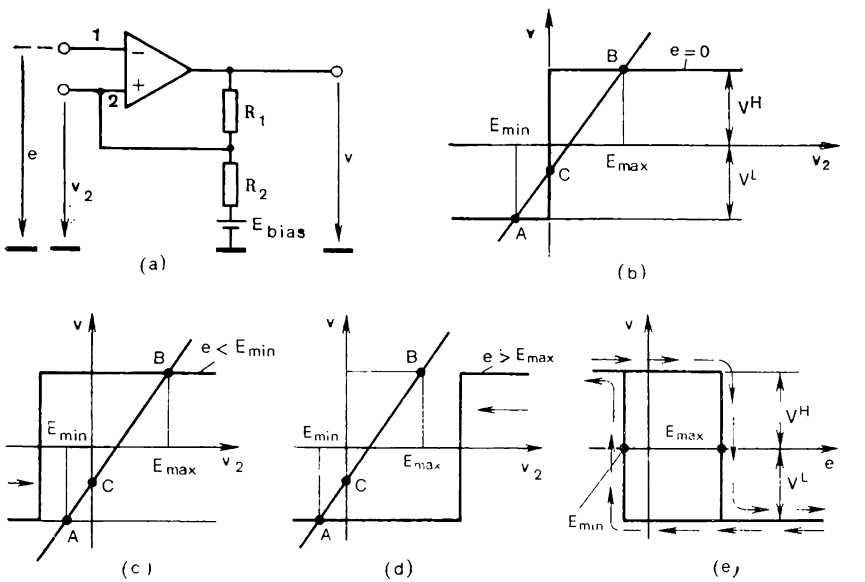


Fig. 6.5

the relay response (V^L and V^H, respectively) with the load line

$$v = -(1-\gamma)\,E_{bias}/\gamma + v_2/\gamma \tag{6.7}$$

and given by

$$\begin{aligned} E_{max} &= \gamma V^H + (1-\gamma)\,E_{bias} \\ E_{min} &= -\gamma V^L + (1-\gamma)\,E_{bias} \end{aligned} \tag{6.8}$$

It follows from Eq. (6.8) that the two thresholds can be shifted at will by varying E_{bias}.

When the voltage at the inverting input is $e < E_{min}$ or $e > E_{max}$, only one point of stable equilibrium remains—point A or B (Fig. 6.5*c* and *d*). Thus, the transfer characteristic $v = f(e)$ has a hysteresis loop (Fig. 6.5*e*), so that the circuit of Fig. 6.5*a* may be regarded as an unbalanced bistable multivibrator. In contrast to the transfer characteristic in Fig. 6.2, an unbalanced bistable multivibrator built around a differential opamp goes from the high state, V^H, to the low state, V^L, as e increases, and vice versa. Therefore, such a circuit must be regarded as an inverter.

6.3. SYMMETRICAL BISTABLE MULTIVIBRATORS

In symmetrical bistable multivibrators, feedback is obtained by connecting two common-emitter inverting switches in cascade (see Fig. 5.15). When the feedback loop from the output of the second

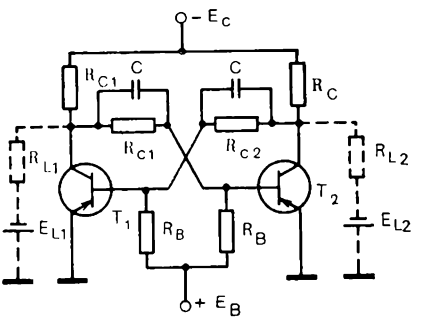


Fig. 6.6

stage to the input of the first is completed, the resultant configuration is a symmetrical bistable multivibrator (Fig. 6.6). In such a circuit, one transistor is normally cutoff and the other is saturated. Assume that the state in which $T2$ is OFF is the 0 state, and the state in which $T1$ is OFF is the 1 state. Because $T1$ is saturated, a balanced bistable multivibrator is likewise called *saturated.*

Operability Condition

As already noted in Sec. 6.1, the conditions for the existence of two stable states in a multivibrator based on logic gates are the same as the conditions for logic gates to reproduce logic levels at their outputs. For the case in question, these conditions in the absence of loading and for $M = N = 1$ are given by Eqs. (5.19) and (5.20). It is also to be noted that symmetrical multivibrators can be optimized in the same manner as advised in Sec. 5.3 for logic gates.

Consider the effect of loading on the static characteristics of a symmetrical bistable multivibrator. In the general case (Fig. 6.6), the load may be unbalanced both in terms of load resistors ($R_{L1} \neq R_{L2}$) and in terms of the potential at the lower terminals of these resistors ($E_{L1} \neq E_{L2}$). For a loaded symmetrical bistable multivibrator to retain operability, it will suffice to satisfy the condition for the transistors to be at saturation. The point is that if the ON transistor (say, $T1$) is saturated, the condition for the second stage ($T2$) to be driven to cutoff will remain unchanged.

Consider the case where R_{L1} is connected in parallel with the collector load resistor R_{C1} and R_{L2} is connected in parallel with $T2$ ($E_{L1} = E_C$ and $E_{L2} = 0$). With this arrangement, both resistors make it more difficult for $T1$ to move into saturation. The effect of R_{L1} is that R_{C1} will now carry only a fraction of the collector current of $T1$, which is equivalent to reducing the current gain of that transistor: $\beta_{eq} = \beta/(1 + R_{C1}/R_{L1})$. The effect of R_{L2} is to reduce the collector voltage, $|V_{C2}| = V^H$, of the OFF transistor $T2$. Neglecting I_{CBO} and e_{EO} because of their smallness, we may write

$$V^H \approx E_C\gamma_c/(\gamma_c + 1 + R_c/R_{L2}) \tag{6.9}$$

From a comparison of Eqs (6.9) and (5.7), it is seen that the effect of R_{L2} is equivalent to raising the loading factor: $N_{eq} = 1 + R_c/R_{L2}$.

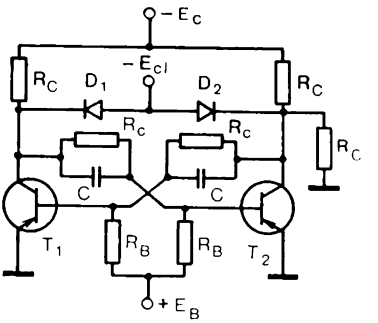


Fig. 6.7

The conditions for a loaded bistable multivibrator to remain operable are defined by Eqs. (5.19) and (5.20) on substituting $\beta_{eq,min}$ and $N_{eq,max}$.

The value of V^H is often stabilized by including clamping diodes (see Fig. 6.7). The collector supply E_C to which the anodes of the clamping diodes are connected is chosen such that the diode connected to the OFF transistor is conducting at a minimum value of R_L. Then, irrespective of loading, $V^H \approx E_{cl}$.

Transient Response

Symmetrical bistable multivibrators can be triggered between the two stable states in any one of several ways. For one thing, the triggering may be *asymmetrical* or *symmetrical.* In asymmetrical triggering, two trigger inputs are employed, one to set the circuit in one stable (say, 1) state, and the other to reset it to the opposite (that is, 0) state. This process is sometimes referred as *set-reset triggering.* Symmetrical triggering uses only one trigger input, and the state of the circuit is changed each time a trigger pulse is applied. There are also bistable multivibrators which use a combination of asymmetrical and symmetrical triggering. They have two inputs for asymmetrical triggering and one more input for symmetrical triggering.

Asymmetrical base triggering. For simplicity, diagram in Fig. 6.8 only shows the triggering-circuit components which serve to switch

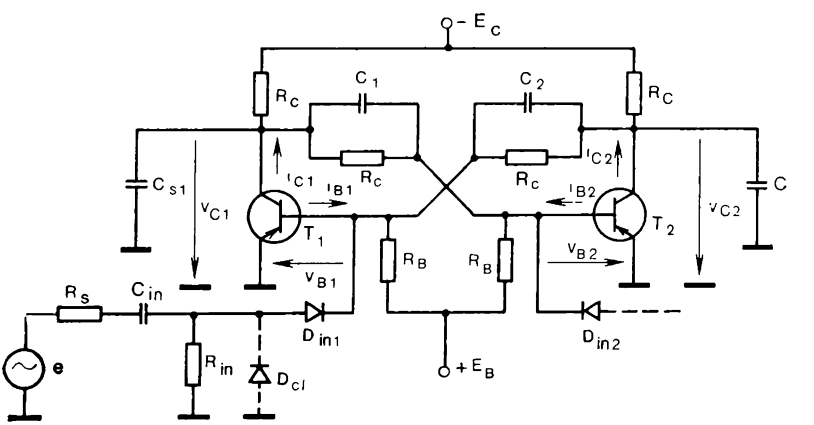


Fig. 6.8

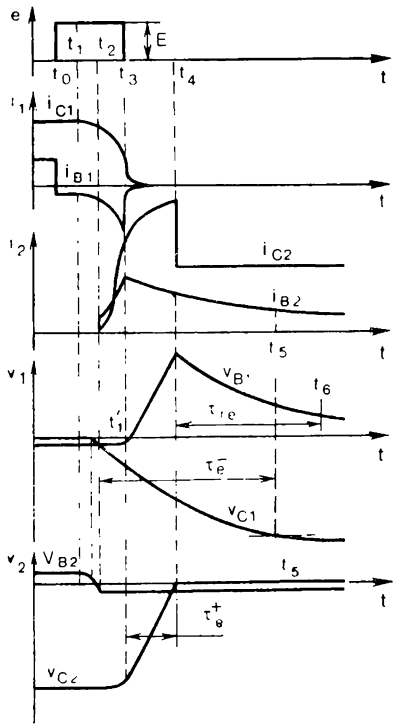


Fig. 6.9

the circuit from the 0 to the 1 state. Switching is done by a positive-going trigger pulse e applied to the base of the saturated transistor $T1$. To avoid interference between the triggering-pulse source and the bistable multivibrator, resort is made to a triggering diode which accepts trigger pulses via a decoupling network, $C_{\text{in}}R_{\text{in}}$. Between trigger pulses, the diode is cut off. As a way of stabilizing the initial voltage, the output of the decoupling network is sometimes coupled to a clamping diode, D_{cl}, in which case R_{d} may be omitted from the circuit.

We shall analyze the transient response of the circuit on assuming that:

—the capacitance of the decoupling capacitor is so high that it passes input pulses without distortion;

—the input resistance of the ON transistor is a small fraction of the internal resistance of the signal source ($r_{\text{in}} \ll R_s + r_t$), so that the signal generator operates under conditions of a short-circuit and generates a current pulse $I_{\text{g}} = E_{\text{g}}/(R_s + r_t)$;

—the multivibrator is switched by a small signal. The conditions under which the signal may be assumed to be a small one will be defined later.

As follows from Fig. 6.9, the total switching time of the bistable multivibrator is the sum of five terms. Within the interval from t_0 to t_1, the trigger pulse applied to the base is followed by a turn-off current $I_{\text{BO}} = I_{\text{g}} - I_{\text{B,ss}}$, where $I_{\text{B,ss}}$, $\approx E_{\text{C}}/R_{\text{C}}\,(1 + \gamma_{\text{c}}) - I_{\text{bias}}$ is the steady-state base current of the saturated transistor. The duration of the interval from t_0 to t_1 is given by Moll's equation, Eq. (3.52), in which $I_1 + I_2 = I_{\text{g}}$ and $K_{\text{sat}} = \beta I_{\text{B,ss}}/I_{\text{C(sat)}}$.

Within the interval from t_1 to t_2, the potentials at the collector of $T1$ and the base of $T2$ keep falling until $T2$ turns on at time t_2. Because the speed-up capacitors are sufficiently large, $C2$ will pass practically all of the change of potential from the collector of $T1$ to the base of $T2$.

If we include the stray capacitance C_{s1}, it may be deemed that the collector potential v_{C1} decreases exponentially, with the initial linear part given by

$$v_{C1\sim}(t) \approx -R_{C1}(I_{C(sat)}+\beta I_{BO})(t-t_1')/\tau_\beta^*$$

where $t_1' - t_1 = \tau_2^*$; for τ_β^*, see Eqs. 3.44-3.46.

Assuming that $T2$ turns on at $v_{B_2} \approx 0$ and the initial cutoff voltage is $V_{B,bias}$, we get

$$\Delta t = t_3 - t_2 = \tau_2^* + V_{B,\,bias}\tau_\beta^*\,[R_C(I_{C(sat)}+\beta I_{BO})]^{-1} \qquad (6.10)$$

In deriving Eq. (6.10), it is assumed that $T2$ turns ON before $T1$ turns OFF, which is true at a low value of I_{BO}. Thus, refining the definition of a small signal, we define it as such that $T2$ turns ON before $T1$ turns OFF.

Within the intervals from t_2 to t_3, both transistors are in the active region, so the feedback loop gain exceeds unity, thereby giving rise to a regenerative process leading to a cumulative decrease in the currents of $T1$ and a cumulative increase in the currents of $T2$. The regeneration interval terminates in driving $T1$ to cutoff. To evaluate the regeneration time, τ_{reg}, we shall present the bistable multivibrator as a two-stage current amplifier (Fig. 6.10). Neglecting the change of voltage on the speed-up capacitors and assuming that the input resistance of the ON transistor is a small fraction of R_C, the Laplace transform of the stage current gain may be written

$$K_i(s) = -\beta/(1+s\tau_\beta) \approx -1/s\tau_B \quad \text{for} \quad \tau_{reg} \ll \tau_\beta$$

where $\tau_B = \tau_\beta/\beta$ is equal to the average transit time for minority carriers in the base.

As is seen from Fig. 6.10, the Laplace transform of the incremental input current is

$$i_{B1\sim}(s) = \frac{I_{in\sim}(s)}{1-K_i^2(s)} \approx \frac{I_{in\sim}(s)}{1-(s^2\tau_B^2)^{-1}}$$

Since at the initial current $i_{B_1} = -i_{C1}(t_2)/\beta$ the system is in a state of equilibrium, we may deem that

$$I_{in\sim} \approx I_{BO} + i_{C1}(t_2)/\beta$$

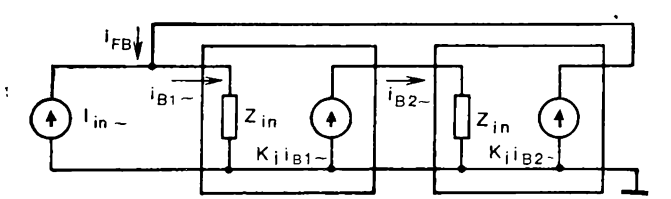


Fig. 6.10

so that

$$i_{B1\sim}(s) \approx \frac{I_{BO} + i_{C1}(t_2)/\beta s}{1 - (s^2\tau_B^2)^{-1}}$$

The Laplace transform of the emitter-current alternating component and the respective original time function are

$$i_{E\sim}(s) = [K_t(s) + 1]\, i_{B\sim}(s)$$

$$i_{E\sim}(t) = [I_{BO} + i_{C1}(t_2)/\beta] \exp[(t - t_2)/\tau_B]$$

The transistor turns off at $i_{E\sim} \approx i_{C1}(t_2)$, whence

$$\tau_{reg} \approx \tau_B \ln \frac{i_{C1}(t_2)}{I_{BO} + i_{C1}(t_2)/\beta} < \tau_B \ln \frac{I_{C(sat)}}{I_{BO} + I_{C(sat)}/\beta}$$

Thus, the regeneration time is proportional to the mean transit time of minority carriers in the base and decreases with increasing current, I_{BO}. For typical parameter values ($I_{BO} = 0$, $\tau_B = 10$ ns, and $\beta = 50$), τ_{reg} is about 40 ns. This implies that even with a small signal the regeneration process in a bistable multivibrator leads to a fast turn-off of the transistor. The transients taking place during the subsequent stages are similar to those in logic gates under large-signal conditions.

The interval from t_0 to t_3 (Fig. 6.9) is the delay time, that is the delay with which the leading edge of v_{C2} in the switch to become conducting rises to its operating value:

$$\tau_d = \tau_s + \Delta t + \tau_{reg} \tag{6.11}$$

From the above analysis it is seen that the delay decreases with increasing triggering current I_{BO} and increases with increasing saturation factor of the ON transistor and increasing cutoff voltage $V_{B,bias}$ at the base of the OFF transistor.

During the interval from t_3 to t_4, the positive-going leading edge of voltage on the collector of the turning-on stage is shaped. The worst case is illustrated in Fig. 6.9 where the input trigger pulse ceases at the instant when $T1$ turns off, so that the leading edge is solely controlled by the rising collector current of $T2$. The duration of the leading edge in this case, τ_e^+, can be found from Eq. (3.46).

Another important process taking place within the interval in question is that the negative-going leading edge of the exponentially decreasing voltage on the collector of $T1$ is shaped. The time τ_e^- terminates at the instant t_6, when v_{C2} rises to $-0.9V^H$. Now, $\tau_e^- \approx 2.3\tau_C$, where $\tau_C = (R_C \parallel R_c)(C + C_s + C_C)$ is the equivalent time constant of the collector circuit.

During the interval from t_4 to t_6, the voltage on the base of the OFF transistor $T1$ reaches its steady-state value. The duration of this transition interval is controlled by the time required for C_1 to charge up. The transition time, also called the dynamic bias

duration or the resolving time of a bistable multivibrator, is given by

$$\tau_{re} \approx 2.3C \ (R_B \| R_c)$$

The capacitance of the speed-up capacitors should preferably be reduced so as to minimize both τ_e^- and τ_{re}. On the other hand, in deriving relations for Δt, τ_{reg}, and τ_e^+, it has been assumed that the value of C is so high that its effect may be neglected. If C be taken too small, the above time intervals might increase. A more detailed transient analysis (see, for example, [4]) shows that an acceptable balance between the desire to reduce τ_e^- and $\tau_{D,c}$, and to maintain Δt, τ_{reg} and τ_e^+ sufficiently small can be struck by choosing the value of the speed-up capacitor such that it is approximately three or four times the ratio $\tau_\beta^*/R_{C1(2)}$:

$$C \approx 3\tau_\beta^*/R_{C1(2)} \text{ to } 4\tau_\beta^*/R_{C1(2)} \tag{6.12}$$

With the ratio τ_e^-/τ_e^+ ranging from seven to nine, the above value of C practically ensures the minimum attainable value of τ_e^+.

Sensitivity. We shall define the sensitivity of a bistable circuit as the minimum duration τ and amplitude of the trigger-pulse current I_g just sufficient for the circuit to change state reliably. Approximately, it may be deemed that the minimum duration of a trigger pulse should be $\tau_{min} \approx t_0 - t_2$, because at t_2 the cutoff voltage is removed from the base of $T2$ and a process of regeneration commences, which may terminate in the absence of a triggering pulse. In view of Eqs. (3.52) and (6.10) the minimum required charge, Q_{min}, of a triggering pulse, defining the sensitivity of a bistable circuit can be written

$$Q_{min} = (I_g\tau)_{min} \approx \tau_{sat} I_{C(sat)} (K_{sat} - 1) + \tau_\beta^* V_{B,\,bias}/(R_C\beta) \tag{6.13}$$

As follows from Eq. (6.13), the sensitivity of a bistable circuit improves with decreasing degree of saturation of the ON transistor and decreasing bias voltage on the base of the OFF transistor. Also, Eq. (6.13) defines the immunity of a bistable circuit towards interfering pulses produced in the base circuit of the ON transistor. More specifically, an improvement in sensitivity impairs the noise margin of the bistable circuit.

Speed. The resolving time, which is a measure of the speed of an asymmetrically base-triggered bistable multivibrator, is stated in terms of the interval from t_0 to t_5:

$$T_{min} = \tau_d + \tau_e^+ + \tau_{re} \tag{6.14}$$

If $R_c \| R_B$ ranges between 5 and 10 kΩ, the principal term of the switching time is the resolving time, τ_{re}. Typical values of T_{min}, for diffused transistors are from 5 to 7 μs, and for alloyed transistors from 1 to 1.5 μs.

Symmetrical base triggering. A major distinction of symmetrical base triggering (Fig. 6.11) is that the OFF transistor remains off

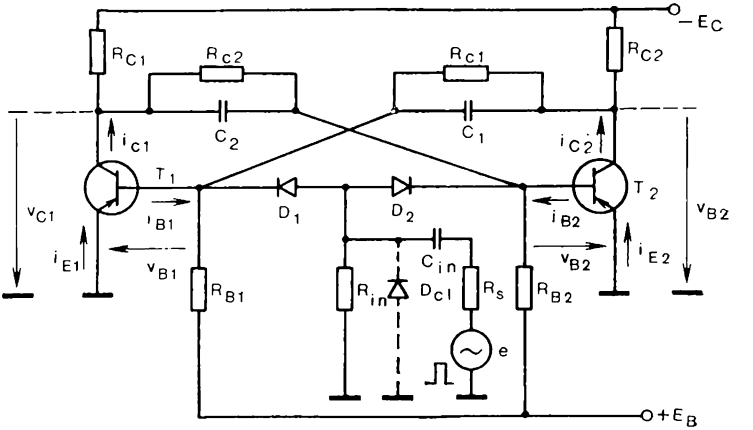


Fig. 6.11a

until the triggering pulse ceases. This is because the triggering diode $D2$ jumps into conduction and clamps the potential on the base of $T2$ to some positive value, so that it cannot turn on until the excess minority-carrier charge is removed from the base of the saturated $T1$ and the base potential of $T2$ begins to fall. Thus, the currents of the previously saturated transistor $T1$ diminish when the feedback loop is open-circuited, and only the relationships applicable to a simple switch (see Chap. 3) hold.

At time t_2, the emitter junction of $T1$ is reverse-biased, and the bases of both transistors are rapidly pulled up to the triggering-pulse voltage. The corresponding positive changes of voltage are passed via the speed-up capacitors to the collectors. Thus, until the triggering pulse ceases at time t_3, transistor $T2$ remains in the OFF state, and only v_{C1} keeps decreasing exponentially during the interval from t_2 to t_3.

At t_3, the cutoff voltage on the transistor bases drops to zero, and there appear positive changes of base current, I_{Bm1} and I_{Bm2}, which bring about an exponential rise in the collector currents. The positive feedback loop between the stages of the bistable circuit is now completed, but the cumulative rise of currents observable in an asymmetrically triggered bistable circuit does not take place because the base currents undergo step changes in the same direction. Within this portion of the transfer characteristic, the speed-up capacitors serve as d.c. blocking capacitors which pass the rising collector currents, thereby reducing the base currents. Obviously, if the initial value of I_{Bm2} is high in comparison with that of I_{Bm1}, the rate of rise of i_{C2} will exceed that of i_{C1}, and i_{B1} will decrease faster than i_{B2}. As a result, i_{C1} will ultimately cease to rise and the

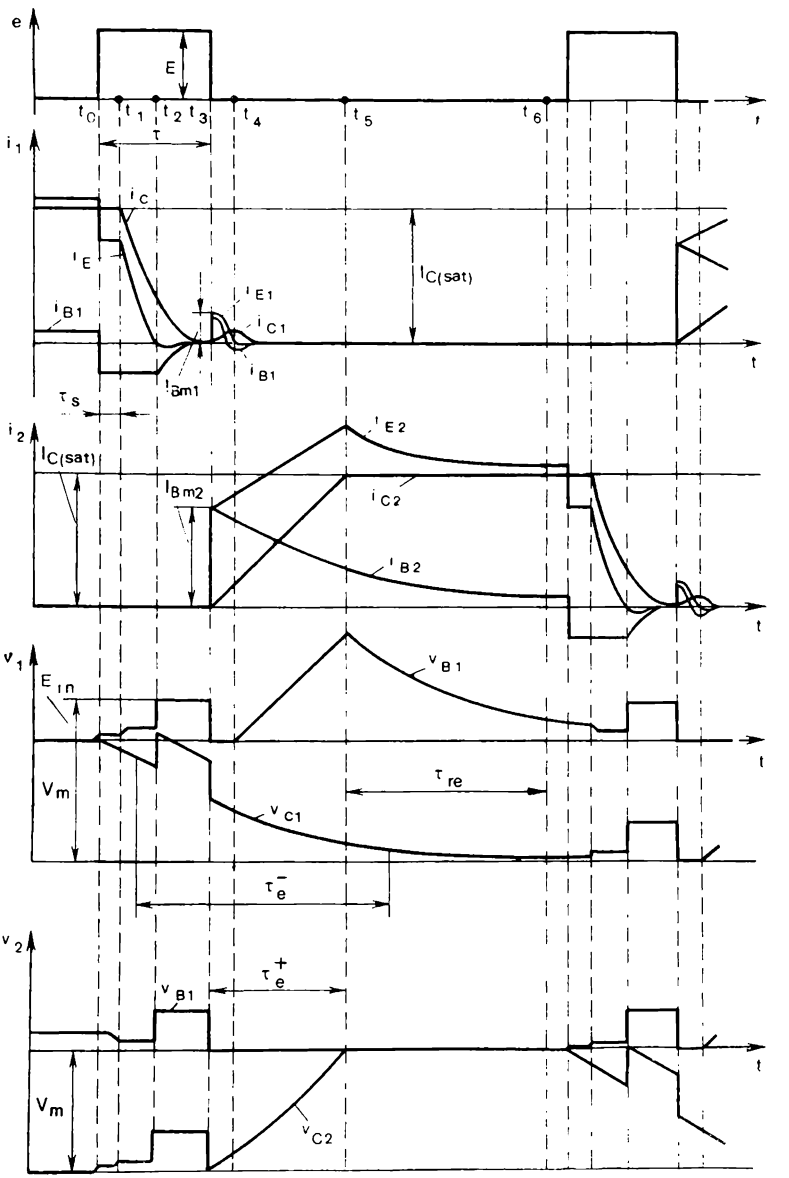


Fig. 6.11b

emitter junction of $T1$ will be reverse-biased (time t_4 in Fig. 6.11*b*). To sum up, the condition for a normal switch-off of $T1$ may be written

$$I_{Bm1} < I_{Bm2} \tag{6.15}$$

According to [4], with perfect symmetry between the arms of the bistable circuit ($R_{C1} = R_{C2}$ and $V_{m1} = V_{m2}$), Eq. (6.15) will be satisfied at any value of the ratio τ/τ_e^-. In real circuits, however, there always is an amount of dissymetry, so that a large absolute difference between I_{Bm1} and I_{Bm2} is essential for the bistable circuit to switch reliably. This requirement is satisfied by increasing the value of speed-up capacitors over their optimal value for an asymmetrically triggered bistable circuit. Assuming a 10-% spread in value between the resistors and the worst possible combination of circuit parameters, a bistable circuit will change state reliably at $\tau/\tau_e^- \leqslant 3$; whence

$$C \geqslant \tau/(3R_C) \tag{6.16}$$

The interval from t_3 to t_4 is usually small in comparison with the duration of the positive and negative leading edges. During the intervals from t_4 to t_5 and from t_5 to t_6, the positive- and negative-going leading edges are formed and the static voltage on the base of the OFF transistor is restored. In general, these processes are analogous to those observed with asymmetrical base triggering. A distinction is that in determining the duration of the negative-going leading edge, one must also include the initial interval from t_1 to t_3 during which the triggering pulse may bring about a substantial change (proportional to the area of the triggering pulse) in the voltage on the collector of the previously ON resistor.

Now the total switching time of a symmetrical base-triggered bistable circuit also includes the duration τ of the triggering pulse. Therefore, for a higher speed, it is desirable to reduce the value of τ by causing it to approach the interval from t_0 to t_2; this may be taken as the minimum acceptable duration of a triggering pulse. As a

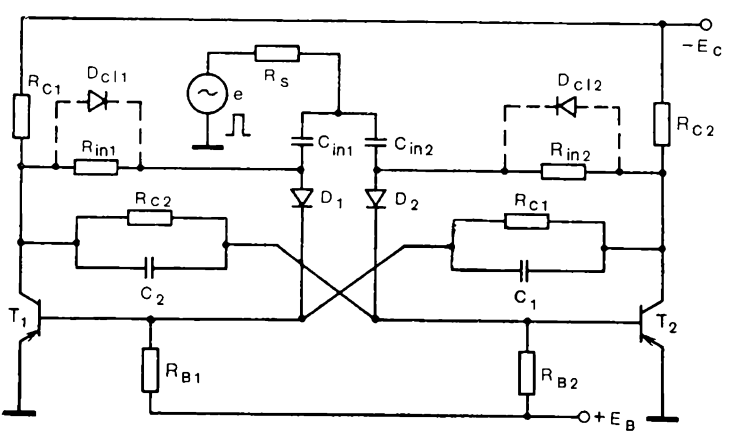


Fig. 6.12

rule, the speed of a symmetrically base-triggered bistable circuit is half as high as that of an asymmetrically base-triggered circuit.

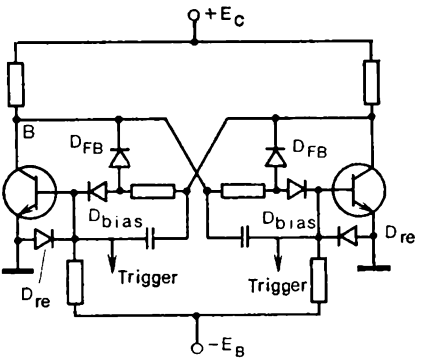


Fig. 6.13

Steered symmetrical base triggering. The speed of a symmetrically base-triggered bistable circuit can be brought closer to that obtained with asymmetrical base triggering by applying the output voltage of the bistable circuit to the anode of a second triggering diode so as to prevent it from jumping into conduction. This is known as *steered triggering.* A circuit adapted to steered symmetrical base triggering is shown in Fig. 6.12.

Unsaturated bistable circuit. The primary factor limiting the speed of bistable circuits using r.f. drift transistors is the storage time, that is, the time required for the removal of the excess minority-carrier charge from the base and collector of a saturated transistor. As is shown in Chap. 3, the storage time can effectively be cut down by applying nonlinear negative feedback. Figure 6.13 shows a high-speed bistable circuit using nonlinear-feedback switches in which the bias voltage is supplied by D_{bias}. Because the transistor in the conducting arm of the bistable circuit is in the active region, this configuration is sometimes referred to as an *unsaturated bistable circuit.*

The resolving time of the circuit in Fig. 6.13 can be cut down by using recovery diodes D_{re} placed across the emitter junctions of the transistors. With this arrangement, the time constant of the circuit over which the charge on the speed-up transistor is restored is Cr_t. The joint use of nonlinear feedback and recovery diodes can cut down the resolving time of a bistable circuit by an order of magnitude.

6.4. VACUUM-TUBE BISTABLE CIRCUITS

Vacuum-tube bistable circuits, both symmetrical and unsymmetrical, are analogous to their transistorized counterparts.

In a symmetrical vacuum-tube bistable circuit (Fig. 6.14), the voltage dividers R_cR_g in the coupling networks are chosen such that in static states one tube is at cutoff and the other is operating with its grid drawing current. With this arrangement, the change of plate voltage will be more stable despite variations in tube parameters and supply voltages. The change of plate voltage is only a fraction

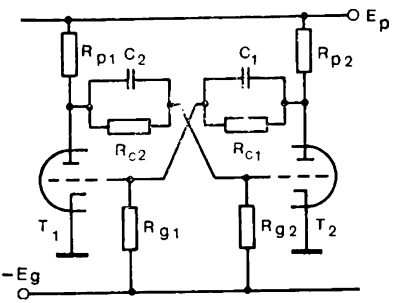


Fig. 6.14

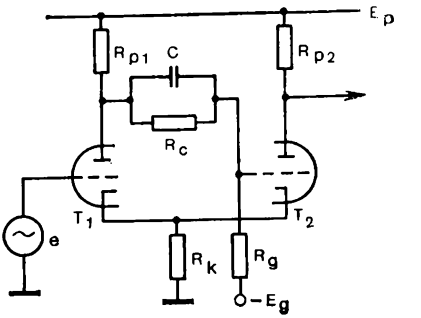


Fig. 6.15

of E_p, because its value for an ON tube is substantially greater than for an ON transistor.

In an unsymmetrical cathode-coupled vacuum-tube bistable circuit (Fig. 6.15), assuming that the circuit is in one of its two stable states ($T1$ is OFF and $T2$ is ON), the initial level of output voltage is stabilized by causing $T2$ to operate with its grid drawing current ($v_{gk2} > 0$). The characteristics of vacuum-tube bistable circuits are given in [4].

6.5. FLIP-FLOPS

When bistable circuits are used in logical digital networks, they are usually called flip-flops. A flip-flop is switched between its two stable states by input signals of the logic type, that is, two voltage levels of which one corresponds to a logic 0 and the other to a logic 1. If the input and output logic levels are properly matched, the output of one flip-flop ($FF1$) can be coupled to the inputs of other flip-flops ($FF2$) and logic gates ($LG2$), and the output of a logic gate ($LG1$) to the input of a flip-flop (Fig. 6.16).

Sequential Circuits and Finite-State Machines

Outwardly, a flip-flop resembles a logic gate, because its output is a single binary variable whose value is determined by the state

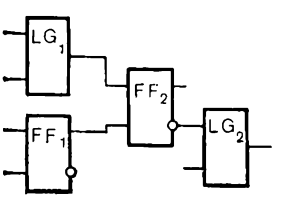


Fig. 6.16

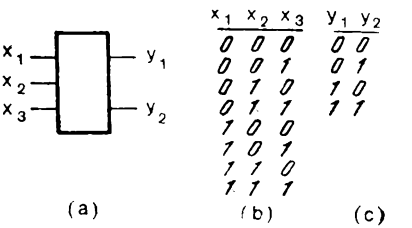


Fig. 6.17

of the flip-flop and the combination of binary variables present at its inputs. In contrast to a logic gate, however, the input-output relations of a flip-flop cannot be described by a switching function. The point is that the existence of two stable states makes the correspondence between input and output other than univalued. Now we must know not only the combinations of the inputs, but also their time sequence. For this reason, such circuits are called *sequential.*

Hence, an important point about sequential circuits (or networks) is that the present value of the output is dependent on the past values of inputs as well as on their current values. For example, the output signal of an asymmetrically triggered flip-flop depends on which of the inputs was the last to accept a trigger pulse. For a symmetrically triggered flip-flop, the output signal depends on whether the input sequence contained an even or an odd number of pulses.

Mathematically, the input-output relations of sequential circuits are modelled by what are known as *finite-state machines* (or still more abstractly, *finite automata*). As abstractions, finite-state machines are convenient tools in a functional analysis of flip-flops and complete digital networks and systems.

Definition. A finite-state machine (a finite automaton) is a system for which the following parameters are defined:

—sampling time t which actually is the ordinal number ($t = 0, 1, 2, \ldots$) of a particular sampling interval;

—a finite set $\{x_1, x_2, \ldots, x_{n1}\}$ of likely values of the input signal $X(t)$, called the input alphabet of the finite-state machine;

—a finite set $\{y_1, y_2, \ldots, y_{n2}\}$ of likely values of the output signal $Y(t)$, called the output alphabet of the finite-state machine;

—a finite set $\{q_1, q_2, \ldots, q_{n3}\}$ of internal states;

—a transition (next-state) function F_t which defines the state into which the finite-state machine will move at time $(t + 1)$, if its state and the value of the input signal at time t are known;

—an output function F_{out} relating the value of the output signal to the internal state of the machine and the input signal at time t.

Denoting the current state of the machine by $Q(t)$, the next-state function F_t and the output function F_{out} may be written

$$Q(t+1) = F_t(Q(t), X(t)) \tag{6.17}$$

$$Y(t) = F_{\text{out}}(Q(t), X(t)) \tag{6.18}$$

If we know the form of F_t and F_{out} for any initial state $Q(0)$ and any input signal $X(t)$, we can find the sequence of states $Q(t)$ and the output signal $Y(t)$.

Sequential circuits most commonly used in the digital field ordinarily have binary inputs and outputs. Accordingly, the input and output alphabets of the finite-state machines modelling them are

sets of combinations of binary input and output variables. Consider as an example a finite-state machine (Fig. 6.17*a*) having three input variables x_1, x_2, and x_3, and two output variables y_1 and y_2. The input alphabet contains eight combinations of variables (Fig. 6.17*b*), the output alphabet four (Fig. 6.17*c*). Sometimes, the input and output alphabets will contain only some of all the likely combinations. The combinations not included are called forbidden.

A fundamental result of the theory of finite-state machines (finite automata theory) [48] is that any sequential circuit can be decomposed into a system of two simpler interacting entities, one having no memory and the other having memory. The memoryless part of a sequential digital (or switching) circuit is what is known as a *combinational* (or *combinatorial*) *logic*. Its present behaviour is determined solely by the present inputs, regardless of the past input history. Thus, a combinational circuit may be thought of as a one-state (trivial) sequential machine. Its next-state function reduces to an identity function, and its output function to a set of switching functions for the input binary variables, that is,

$$\begin{aligned} y_1 &= F_{\text{out1}}(x_1, x_2, \ldots, x_{M1}) \\ y_2 &= F_{\text{out2}}(x_1, x_2, \ldots, x_{M1}) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_{M2} &= F_{\text{out},\,M2}(x_1, x_2, \ldots, x_{M1}) \end{aligned} \tag{6.19}$$

where $M1$ and $M2$ are the number of input and output binary variables respectively.

The memory part of a sequential digital circuit uses discrete storage elements, such as flip-flops. In effect a flip-flop may be looked upon as an elementary sequential circuit, or a finite-state machine with two internal states. Its states are coded by a binary variable Q whose value is the same as that of the output variable, so its output function reduces to an equation

$$y = Q \tag{6.20}$$

Flip-flops differ in the number of inputs and in the form of the next-state (transition) function. Digital networks are usually made with a relatively small number of different flip-flop types. In principle, any sequential circuit can be realized with a single type of flip-flop, provided its transition function statisfies the condition of completeness, which requires that for any state there must be an input signal which will switch the flip-flop to a new state, and an input signal which will leave the state of the flip-flop unaffected.

Before we take up the actual flip-flops that go to make digital sequential circuits most frequently, it appears useful, as we believe, to examine them as finite-state machines. In our discussion, we shall use the same type designations as are assigned to the flip-

flops themselves. The reader interested in the flip-flops as such may skip this discourse and go directly to the flip-flop circuits.

***D*-type elementary finite-state machine.** This form of finite-state machine has one binary input variable, D. The transition (next-state) function has the form

$$Q(t+1) = D(t) \tag{6.21}$$

In other words, the state of the machine at time $(t+1)$ is independent of its previous state and takes the value of the input variable at time t. In digital systems, a D-type machine delays the input signal for one bit time.

***T*-type elementary finite-state machine.** This form of finite-state machine has one binary input variable, T. Its transition (next-state) function is described by a switching function of the form

$$Q(t+1) = \bar{T}(t)\,Q(t) + T(t)\,\bar{Q}(t) = T(t) \oplus Q(t) \tag{6.22}$$

In other words, the output signal at time $(t+1)$ is the modulo 2 sum of the machine state and the input signal at time t. It is denoted by the $\oplus$ sign. The function of a T-type machine is to count input signals one by one.

***RS*-type elementary finite-state machine.** This form of finite state machine has two binary input variables, R and S. Its transition (next-state) function is described by a switching function of the form

$$Q(t+1) = S(t) + Q(t)\,\bar{R}(t) \tag{6.23}$$

provided that $RS = 0$. The machine will retain its internal state at $R = S = 0$. When $S = 1$, the machine is set to a logic 1 state; when $R = 1$, the machine is reset to a logic 0 state. The input condition $S = 1$ and $R = 1$ is an indeterminate one and should never be allowed to occur. The function of an RS-type machine is to delay the input signal for a bit time or to store one bit of data.

***DV*-type elementary finite-state machine.** This form of machine has two binary input variables, D and V. Its transition function is described by a switching function of the form

$$Q(t+1) = V(t)\,D(t) + \overline{V(t)}\,Q(t) \tag{6.24}$$

When $V=1$, the machine is analogous to a D-type machine; when $V = 0$, it retains its state unchanged.

A DV-type machine is universal, that is, it can delay the signal for one bit time, store one bit of data, and count input signals one by one. To perform counting, it will suffice to connect the $\bar{Q}$ output to the D input in a DV-type machine. Then Eq. (6.24) will reduce to Eq. (6.22).

***JK*-type finite-state machine.** This form of finite-state machine has two binary input variables, J and K. Its transition function is

described by a switching function of the form

$$Q(t+1) = J(t)\overline{Q(t)} + \overline{K(t)}\,Q(t) \tag{6.25}$$

This is a universal machine. When $JK = 0$, it operates like an RS-type machine. When $J = K$, the machine operates like a T-type machine.

Accordingly, the physical realizations of the above finite-state machines—flip-flops—can have the following inputs:

S to set a flip-flop to the 1 state
R to reset a flip-flop to the 0 state
T to toggle (or complement) the previous state of a flip-flop
D to set or reset a flip-flop
V to enable a change of state
J to switch from the 0 to the 1 state
K to switch from the 1 to the 0 state

In contrast to abstract finite-state machines, real pulse-output and digital systems operate in continuous time. Conversion from the discrete-time operation associated with abstract finite-state machines to the continuous-time operation of real digital systems is an important division in the theory of automata.

According to the manner in which their operation is timed, all sequential networks, including flip-flops, may be divided into asynchronous and synchronous.

Asynchronous Sequential Networks

These are mechanized with level logic gates. In them, the signals are allowed to propagate through the logic chains at their own pace, because in a normally operating sequential network the logic gates change state only in response to a change in the input signal.

Each change in the input signal is followed by a transient process during which the interconnected logic gates switch in turn. This process terminates in the sequential network assuming a new stable state in which it will remain until the input is changed again. Because level logic gates contain no dynamic storage elements, the time during which the network remains in a stable state does not affect the processes accompanying the subsequent transition. Therefore, it is legitimate to adopt as the unit of time the interval between two consecutive changes in the input. Then at the beginning of each transition, what we call sampling time will take on an integer value for which the state of the sequential network is known.

As an example, let us define sampling time for an asynchronous sequential network with three binary input variables whose timing diagrams are shown in Fig. 6.18*a*. The instants at which transitions take place are numbered from 1 through 6. Taking the interval between two consecutive transitions as the unit of time we may

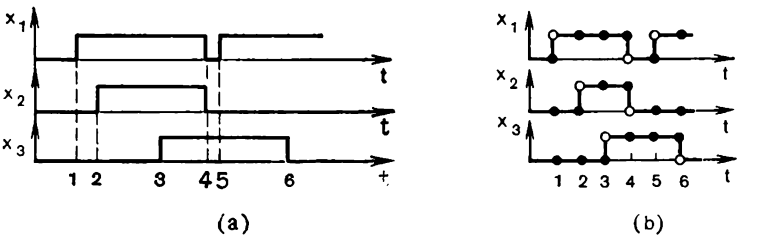


Fig. 6.18

introduce sampling time t' which takes on an integer value (marked by the open and filled circle in Fig. 6.18*b*) when the input variables change in value. At these integer values of t', the switched variable is assumed to have the same value as it had immediately before the transition.

Input alphabet of an asynchronous sequential network. Suppose that we know the combination of the input variables and the internal state of the system at some sampling time t. The next state to which the network will go at time $(t + 1)$ depends on which of the input variables changes its value at time t. Thus, the input signal of an asynchronous sequential network is defined not only in terms of the binary input variables, but also in terms of their change at discrete instants of time.

Let us connect each input variable x to a binary variable dx, a digital differential defined for integer values of t and taking on a value of 1 if x changes and a value of 0 if x does not change at that instant. The value of dx can be represented as a switching function for the values of x at times t and $t + 1$:

$$dx(t) = x(t+1) \oplus x(t) \tag{6.26}$$

In other words, the input alphabet of an asynchronous sequential network is determined by the sum of the combinations of the binary input variables and of their differentials.

The sequence of input combinations for the sequential network illustrated in Fig. 6.18 is given in Table 6.1.

Table 6.1. Sequence of Input Combinations for an Asynchronous Sequential Network

t	x_1	dx_1	x_2	dx_2	x_3	dx_3	t	x_1	dx_1	x_2	dx_2	x_3	dx_3
1	0	1	0	0	0	0	4	1	1	1	1	1	0
2	1	0	0	1	0	0	5	0	1	0	0	1	0
3	1	0	1	0	0	1	6	1	0	0	0	1	1

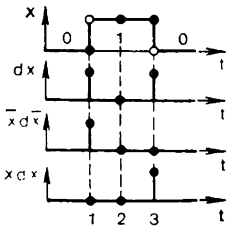


Fig. 6.19

In analyzing the next-state functions of an asynchronous flip-flop, it is required to describe the transitions of the input variables causing changes of state. This can conveniently be done by using the logical (Boolean) product of the variables (or of its negation) by its differential. The unity value of the product $x\,(t)\,\mathrm{d}x\,(t)$ will then indicate that at time t the variable x changes from the 1 to the 0 state. Let us call this transition the negative edge of the signal $x\,(t)$. Similarly, the unity value of the product $\overline{x\,(t)}\,\mathrm{d}x\,(t)$ indicates that at time t the variable x changes from 0 to 1. Let us call this transition the positive edge of the signal $x\,(t)$. Figure 6.19 shows the sequence of signals $\mathrm{d}x$, $x\mathrm{d}x$ and $\bar{x}\mathrm{d}x$ for a positive pulse.

In the general case, the next-state function of an asynchronous flip-flop is the switching function of the flip-flop's internal state Q, input variables $x_1, x_2, \ldots, x_M$ and their differentials:

$$Q\,(t+1) = F_t\,[Q\,(t),\ x_1\,(t),\ \ldots,\ x_M\,(t),\ \mathrm{d}x_1\,(t),\ \ldots,\ \mathrm{d}x_M\,(t)] \tag{6.27a}$$

It can be recast in differential form as

$$\mathrm{d}Q\,(t) = \varphi_t\,[Q\,(t);\ x_1\,(t),\ \ldots,\ x_M\,(t);\ \mathrm{d}x_1\,(t),\ \ldots,\ \mathrm{d}x_M\,(t)] \tag{6.27b}$$

Equations (6.27a) and (6.27b) are connected by an equality of the form

$$Q(t+1) = Q(t) \oplus \mathrm{d}Q(t) \tag{6.28}$$

Suppose that the state of an input of an asynchronous flip-flop is determined by the input variable x and its differential dx. They can affect the transition of the flip-flop to the next state in any one of several ways. Accordingly, there are several types of input.

A flip-flop is said to have a dynamic input, if the transitions of the flip-flop are solely controlled by a change in the value of the input variable x. As a further subdivision, there may be a normal dynamic input, if the drive is the positive-going edge of the signal $x\,(t)$, or a complementing (inverting) dynamic input, if the drive is the negative-going edge of the signal $x\,(t)$. The dependence of the next-state function on the variables $x\,(t)$ and $\mathrm{d}x\,(t)$ existing at a dynamic input can be written as a function of only one binary variable equal to the positive-going edge, $\bar{x}\mathrm{d}x$, in the case of a normal input, or to the negative-going edge, $x\mathrm{d}x$, in the case of a complementing input. The very possibility of such a presentation is an indica-

tion that the input is a dynamic one. Graphic symbols for normal and complementing inputs are shown in Figs. 6.20*a* and *b*.

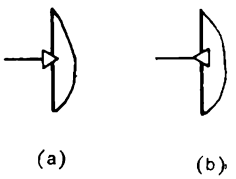


Fig. 6.20

A flip-flop is said to have a static input, if the value of the input variable affects its transitions, even if the value remains unchanged during a transition. As a further subdivision, there may be a steering (or control) static input, if a change in x only determines how the other inputs can affect the flip-flop's transitions. An indication of a steering input is the fact that the next-state function of the flip-flop is independent of the differential of the variable, dx, existing at that input.

The inputs of a flip-flop are termed data inputs, if the permitted combinations of input variables $x_1, x_2, \ldots, x_M$ can be identified with any one of three types. Input combinations of the first type (let us call them passive) are those with which the flip-flop retains the property of bistable operation. A change of state may occur due to signals at other (non-data) inputs. Input combinations of the second type (let us call them *reset* or *clear* signals) are those which override the bistable operation of the flip-flop and force it to move to the 0 state ($Q = 0$), irrespective of the flip-flop's previous state and signals at the other inputs. Input combinations of the third type (let us call them *set* signals) are those which set the flip-flop to the 1 state, irrespective of the flip-flop's previous state and signals at other inputs. Combinations which do not fall in any of the three types are forbidden.

Let us define the switching functions $r(x_1, \ldots, x_M)$ and $s(x_1, \ldots, x_M)$ under the following assumptions:

—the r-function takes on value 1, if the combination of $x_1, \ldots, x_M$ is a reset or clear signal, or value 0 otherwise. Let us call it the reset or clear function;

—the s-function takes on value 1, if the combination of $x_1, \ldots, x_M$ is a set signal, and value 0 otherwise. Let us call it the set function.

It follows from the foregoing that

$$rs = 0 \tag{6.29}$$

So long as the signals remain unchanged at a flip-flop's inputs other than data inputs, its state is described as

$$Q = s + Qr \tag{6.30}$$

The validity of Eq. (6.30) can readily be verified. At $s = 1$, $Q = 1$. At $r = 1$, it follows from Eq. (6.29) that $s = 0$; therefore, $Q = 0$. When $r = s = 0$, Eq. (6.30) reduces to $Q = Q$ which has two solu-

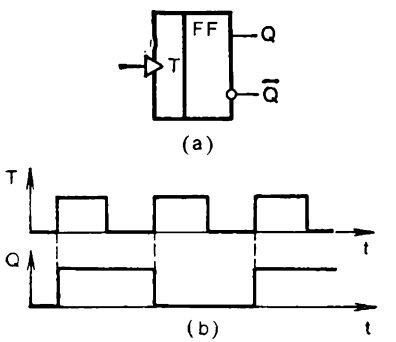


Fig. 6.21

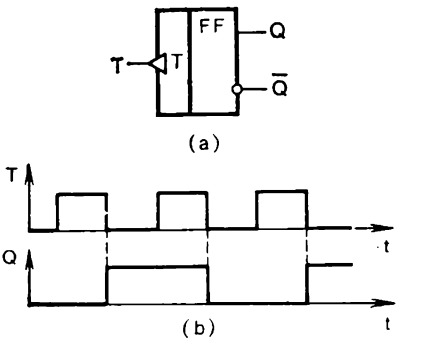


Fig. 6.22

tions: $Q = 0$ and $Q = 1$, which is an indication that the flip-flop retains its property of bistable operation.

The next-state function of a flip-flop having data inputs is defined by

$$dQ(t) = \bar{Q}(t)\,\overline{s(t)}\,ds(t) + Q(t)\,\overline{r(t)}\,dr(t) \tag{6.31}$$

$$Q(t+1) = \overline{s(t)}\,ds(t) + Q(t)\,\overline{\overline{r(t)}\,dr(t)} \tag{6.32}$$

Using Eq. (6.28), we can readily pass from Eq. (6.31) to Eq. (6.32) on the basis of formal transformations in Boolean algebra.

Let us define the next-state functions for the basic types of asynchronous flip-flops implementing the elementary finite-state machines examined earlier.

Asynchronous *T* flip-flop. Let the logic 1 at the T input of the finite-state machine correspond to a positive step change in the input signal $T(t)$ of an asynchronous T flip-flop. Then the next-state function defined by Eq. (6.22) will change to a next-state function of the form

$$Q(t+1) = Q(t) \oplus \bar{T}\,dT \tag{6.33a}$$

or

$$dQ(t) = \bar{T}dT \tag{6.33b}$$

The T flip-flop described by the next-state functions defined by Eqs. (6.33) is triggered by the positive edge of the input signal. Its logic symbol and timing diagrams are given in Fig. 6.21*a* and *b*, respectively. Figure 6.22 shows the logic symbol and timing diagrams for the T flip-flop triggered by the negative-going edge of the trigger input.

Asynchronous *RS* flip-flop. The R and S inputs are data inputs. The set and reset functions are $s = \bar{R}S$ and $r = \bar{S}R$. Since the condi-

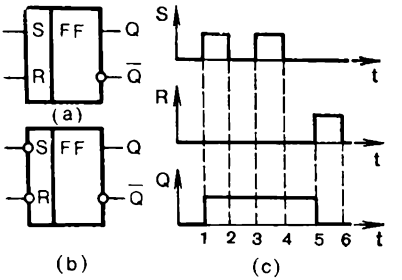


Fig. 6.23

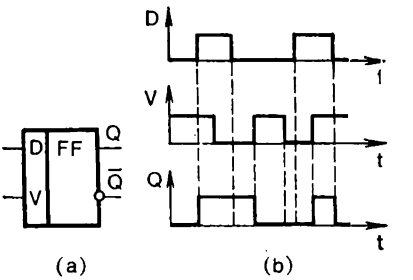


Fig. 6.24

tion $RS = 1$ is never allowed to occur, the set and reset functions may be simplified: $s = S$ and $r = R$. Then the state of the flip-flop may be written

$$Q = S + Q\bar{R} \tag{6.34}$$

The logic symbol for the RS flip-flop is shown in Fig. 6.23*a*. A modification of the above RS flip-flop is one with complementing data inputs. Its logic symbol appears in Fig. 6.23*b*, and its set and reset functions are $s = \bar{S}$ and $r = \bar{R}$.

Asynchronous *DV* flip-flop with set-reset inputs. The set and reset functions are

$$s = VD, \quad r = V\bar{D} \tag{6.35}$$

and the state of the flip-flop may be written

$$Q = VD + \overline{QV}\bar{D} \tag{6.36}$$

The next-state function in differential form is

$$\begin{aligned} \mathrm{d}Q(t) = {} & \bar{Q}(t)\,\overline{V(t)\,D(t)}\,\mathrm{d}\,[V(t)\,D(t)] \\ & + Q(t)\,\overline{V(t)\,\bar{D}(t)}\,\mathrm{d}\,[V(t)\,\bar{D}(t)] \end{aligned} \tag{6.37}$$

The logic symbol and timing diagrams of the DV flip-flop are given in Fig. 6.24*a* and *b*.

Asynchronous *DV* flip-flop with a normal dynamic *V* input and a steering *D* input. The next-state function for this flip-flop is derived from Eq. (6.24) on replacing the variable V with its dynamic equivalent $\bar{V}\mathrm{d}V$. Then,

$$Q(t+1) = D(t)\,\overline{V(t)}\,\mathrm{d}V(t) + Q\overline{\overline{V(t)}\,\mathrm{d}V(t)} \tag{6.38a}$$

or

$$\mathrm{d}Q(t) = [D(t) \oplus Q(t)]\,\overline{V(t)}\,\mathrm{d}V(t) \tag{6.38b}$$

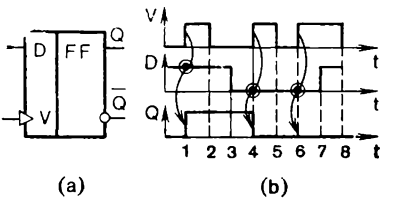


Fig. 6.25

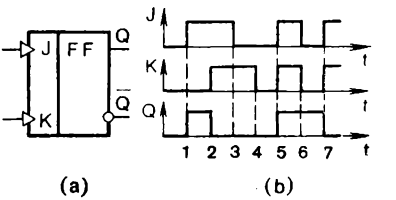


Fig. 6.26

The logic symbol and timing diagrams for this flip-flop appear in Fig. 6.25*a* and *b*. An asynchronous *DV* flip-flop with the complementing *D* or *V* input can be defined in a similar way.

Asynchronous *JK* flip-flop. This flip-flop has two dynamic inputs. The logic symbol for a *JK* flip-flop with normal dynamic inputs is shown in Fig. 6.26*a*. Its next-state functions can be derived from Eq. (6.25) on replacing the variables J and K with their dynamic equivalents $\overline{J}\mathrm{d}J$ and $\overline{K}\mathrm{d}K$. Then,

$$Q(t+1) = \overline{Q}(t)\,\overline{J(t)}\,\mathrm{d}J(t) + Q(t)\,\overline{\overline{K(t)}\,\mathrm{d}K(t)} \qquad (6.39a)$$

or

$$\mathrm{d}Q(t) = \overline{Q(t)\,J(t)}\,\mathrm{d}J(t) + Q(t)\,\overline{K(t)}\,\mathrm{d}K(t) \qquad (6.39b)$$

The timing diagrams for an asynchronous *JK* flip-flop with normal inputs are shown in Fig. 6.26*b*. An asynchronous *JK* flip-flop with complementing dynamic inputs can be defined in a similar way.

Synchronous Sequential Networks

Transition of synchronous sequential circuits from one state at a sampling time t to the other state at the next sampling time $t + 1$ are controlled (synchronized) by what are known as *clock pulses*. Clock pulses are generated by a circuit called the *clock*, and applied to a separate "clock" input. Synchronous sequential circuits may have one or several clock inputs.

The state of a complex sequential circuit is determined by the states of all the flip-flops it contains. Therefore, a transition of a synchronous sequential circuit to a new state is in effect a parallel or a series-parallel transition of the flip-flops in response to clock pulses. Flip-flops which change state under the action of clock pulses are called *clocked* or *synchronous*; they may likewise have one or several clock inputs.

Clocked flip-flops may be classed into two broad groups. In those belonging to the first group, no restrictions are imposed on interconnection between flip-flops; for example, the input of such a flip-

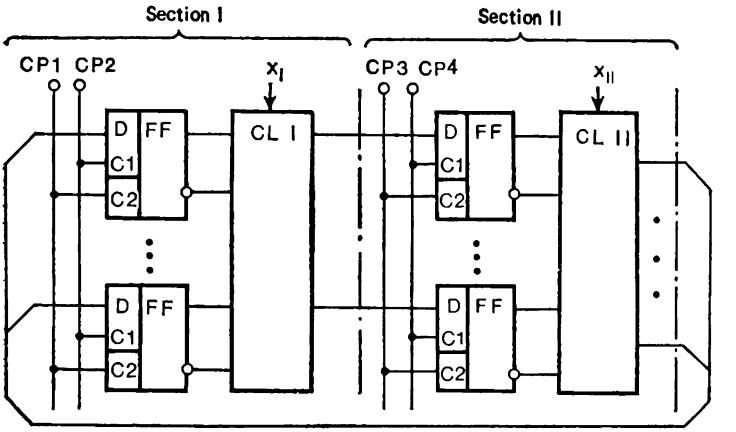


Fig. 6.27

flop may be coupled to its output, the same clock pulses can be used for the component flip-flops and for the sequential circuit as a whole, and the number of clock inputs will be the same for the individual flip-flops and for the entire sequential circuit. In those belonging to the second group, interconnections between the flip-flops within a sequential circuit are allowed only if changes of state occur at different clock times; an interconnection between the input and output of such a flip-flop is not allowed. In our further discussion, flip-flops in the first group will be referred to as having internal (inherent) delay, and those in the second, as having no internal delay, or primitive.

The manner in which primitive flip-flops may be interconnected can conveniently be examined by reference to the diagram of a generalized sequential circuit (Fig. 6.27). It uses D flip-flops which require two clock pulses to change state, so they switch from one state to the other in two steps. During the first step, a clock pulse applied to input $C2$ causes the flip-flop to go to the 0 state, $Q = 0$. During the second step, a second clock pulse applied to input $C1$ causes the flip-flop to switch to the 1 state, $Q = 1$, only if a 1 is applied to the D input. If a 0 is applied to the D input, the flip-flop will remain in the 0 state, $Q = 0$. For the correct operation of the flip-flop, it is essential that the signal at the D input should remain constant so long as there is a pulse at input $C1$. Because of this restriction, flip-flops driven by a common clock pulse may not be interconnected.

Arbitrarily, the sequential circuit of Fig. 6.27 may be divided into two interacting sections. The flip-flops in section I are driven by clock pulses $CP1$ and $CP2$ during the time interval τ_1 (Fig. 6.28), and the flip-flops in section II are triggered by clock pulses $CP3$

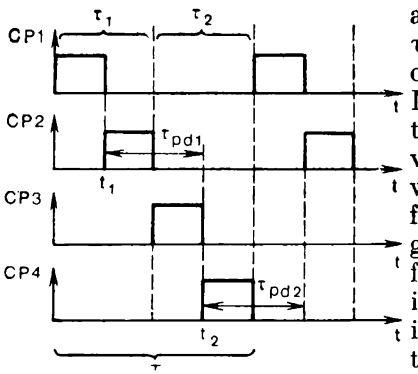


Fig. 6.28

and *CP4* during the time interval τ_2. The two time intervals do not overlap, so the flip-flops in section II remain in their previous states as those in section I change state, and vice versa. With this arrangement, we may interconnect any two flip-flops in the two sections. In the general case, the desired next-state function is obtained by forming the input data signals of the flip-flops in the section as the switching functions of the output signals from the flip-flops in the other. These functions are generatedby combinational logic circuits, *CL1* and *CL2*.

Because it takes some time for a signal to propagate through logic gates, the logic signal arrives at the input of the respective combinational circuit with a delay relative to the instant when a clock pulse is applied to the flip-flops in a given section. For correct operation it is required that all transients at the output of *CL1* should die out before *CP4* is applied. In other words, the maximum propagation delay time through the logic gates in section *I* should not exceed the time interval between the application of *CP2* and *CP4* (the time interval τ_{pd1} in Fig. 6.28). Similarly, the maximum propagation delay through the gates in section II should not exceed τ_{pd2}. These

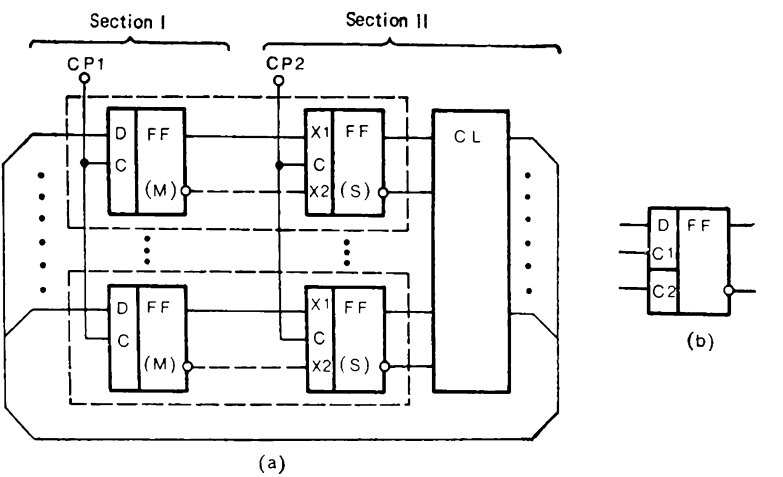


Fig. 6.29

restrictions define the minimum clock period T, equal to the time interval required for the sequential circuit to go from the present to the next state.

Sequential circuits implemented with one-clock-input flip-flops without internal delay are built along similar lines, but they require only two clock pulses for a change of state.

A special case is the sequential circuit in which $CL1$ is omitted (Fig. 6.29*a*). The flip-flops in section I are directly coupled to those in section II. The coupling lines (shown by the dashed lines in the diagram) may be single-wire or two-wire. Two directly coupled basic flip-flops may be regarded as a single flip-flop with an internal delay. The behaviour of this combined flip-flop is determined by that in section I, called the *master*; accordingly, the flip-flop in section II is called the *slave*. Together, they form a *master-slave* (*MS*) *flip-flop*. The logic symbol for a master-slave flip-flop with two clock inputs is shown in Fig. 6.29*b*. Similarly, we can build master-slave flip-flops with three and four clock inputs.

Simulation of Synchronous Flip-Flops by Asynchronous Flip-Flops

A distinction of flip-flops built from level-type logic gates is that their clock inputs can serve as their data inputs. Moreorver, the sequence of clock pulses may be regarded as an alternating sequence of 0 and 1 values of the variable C at the clock input. Because of this, it is legitimate to treat an isolated synchronous flip-flop as an asynchronous circuit in which some of the data inputs serve as clock inputs. In effect, such an asynchronous circuit simulates a synchronous flip-flop in a clocked network.

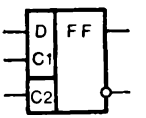
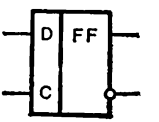
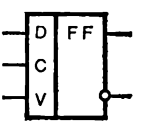
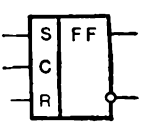
Primitive synchronous flip-flops can be simulated by asynchronous circuits and set-reset inputs because the data input signals of a primitive flip-flop remain constant for the duration of a clock pulse. The state to which a primitive flip-flop can go is uniquely decided by the combination of the input variables. For a primitive flip-flop with a direct clock input C and data inputs $x_1, \ldots, x_M$, the set (1) and reset (0) functions can be written

$$s = C\varphi_1(x_1, \ldots, x_M), \quad r = C\varphi_0(x_1, \ldots, x_M) \tag{6.40}$$

where the form of the functions φ_1 and φ_0 depends on the specific type of flip-flop.

The switching functions and logic symbols for the basic types of direct-input primitive flip-flops are summarized in Table 6.2. For a complementing-input primitive flip-flop, the variable C in Eqs. (6.40) and the respective expressions in Table 6.2 must be replaced by their negations.

Table 6.2. Basic Types of Primitive Flip-Flops

Name	Logic symbol	Switching function
D flip-flop: with two clock inputs		$s=C2D$ $r=\overline{C1}$ $Q=C2D+QC1$
with one clock input		$s=CD$ $r=C\overline{D}$ $Q=DC+Q\overline{\overline{D}C}$
With one clock input: *DV* flip-flop		$s=CDV$ $r=C\overline{D}V$ $\overline{Q}=DVC+Q\overline{\overline{D}VC}$
RS flip-flop		$s=CS$ $r=CR$ $Q=SC+Q\overline{RC}$

Consider the simulation of one-clock-input flip-flops with an internal delay, assuming that in the general case this type of flip-flop implements an elementary finite-state machine for which the next-state function is

$$Q(t+1)=F_t[Q(t),\ X(t)]$$

where $X(t)$ is the binary variables at the data inputs of the flip-flop at time t.

Let Q and X be sampled at the instants when the positive-going edge of a clock pulse C is formed. Then the circuit simulating the flip-flop in question may be visualized as an asynchronous flip-flop which changes state at the instant when the positive-going edge of $C(t)$ is formed, on the condition that $Q \neq F_t(Q,\ X)$. The applicable next-state function has the form

$$dQ(t)=[Q(t) \oplus F_t(Q(t),\ X(t))]\,\overline{C}\,dC$$

or

$$Q(t+1)=F_t[Q(t),\ X(t)]\,\overline{C}\,dC+Q\,\overline{(\overline{C}\,dC)} \tag{6.41}$$

As is seen, the simulated flip-flop has a direct dynamic input C and control inputs X.

In a similar way, we can simulate a synchronous flip-flop in which the values of the data signals Q and X are determined at the instant when the negative-going edge of a clock pulse is formed. To this end, it will suffice to replace the direct dynamic C input in the model by a complementing input.

As an example, consider an asynchronous model of a one-clock-input synchronous D flip-flop. On setting $F_t = D$ in Eq. (6.41), we get

$$dQ = [Q(t) \oplus D(t)]\,\overline{C}\,dC \tag{6.42a}$$

$$Q(t+1) = D(t)\,\overline{C}\,dC + Q(t)\,\overline{(\overline{C}\,dC)} \tag{6.42b}$$

From a comparison with Eq. (6.38) it is seen that a synchronous D flip-flop can well be simulated by an asynchronous DV flip-flop with a dynamic V input.

An advantage of synchronous flip-flops with a dynamic clock input is that there is no restriction imposed on the manner in which signals at its data inputs may change. A disadvantage is that some types of elementary finite-state machines are difficult to realize.

In some cases, a flip-flop circuit can be simplified by imposing reasonable restrictions on changes in the input data signals. The rationale of such restrictions is to forbid any changes in these signals during the half-cycle of a clock pulse preceding the edge that causes the flip-flop to change state. To make the matter more specific, suppose that a flip-flop is caused to change state by the positive-going edge of a clock pulse $C(t)$. Let us require that the time τ_1 during which $C = 1$ be longer than the signal propagation delay through the combinational logic which couples the outputs of the flip-flops to their inputs. Then all transients in the circuit will have died out by the time the negative-going edge of the clock pulse forms, and the binary variables at the data inputs of the flip-flop will remain unchanged for the time interval τ_0 when $C = 0$. In fact, the data variables $Q(t)$ and $X(t)$ of the sequential circuit are sampled during the time intervals τ_1. Subject to the above restrictions, the next-state function for the flip-flop may be written

$$\overline{C}(t)\,dX(t) = 0 \tag{6.43}$$

The equation describing the restrictions on the input contains the variable C without a differential, so the C input must be regarded as the complementing static clock input of a flip-flop with an internal delay. A flip-flop with a normal clock input can be described in a similar way, except that the variable C must be negated in Eqs. (6.41) and (6.42).

6.6. FLIP-FLOPS MADE WITH LEVEL LOGIC GATES

Flip-flops are realized with logic gates which produce the basic switching functions, namely NAND, NOR, AND-NOR, and OR-NAND. Owing to the duality of logic (Boolean) multiplication and addition (see Chap. 5), any flip-flop built from NAND (AND-NOR) gates can be replaced by one made with NOR (OR-NAND) gates without changing the interconnections between the gates. The flip-flop thus obtained will only differ in that it will negate the input and output variables. In view of the foregoing, we may limit our discussion to flip-flops consisting of NAND and AND-NOR gates.

In flip-flops, it is convenient to visualize the NAND gates as a circuit which can be disabled (blocked) and enabled (unblocked) by a control signal. The control signal may be any of the input binary variables. As an example, consider a three-input NAND gate (Fig. 6.30).

Suppose that the control signal is the variable x_1. The gate is disabled (blocked) when $x_1 = 0$, because the state of the output variable, $y = 1$, is then independent of the state of the input variables $x_2, \ldots, x_n$. If $x_1 = 1$, then $y = \overline{x_2, \ldots, x_n}$, which is another way of saying that the output variable depends on the state of the other inputs. As an alternative, we may use as a control signal two or more input variables simultaneously. The gate will then be disabled (blocked) if at least one of the control inputs is in the 0 state, and enabled (unblocked), if a 1 is applied to all the inputs.

When building flip-flops with AND-NOR gates, the function of a unidirectional circuit may be performed by using a **distributed AND connection.**

1-bit storage cell. As is shown in Sec. 6.1, a circuit made up of two inverters connected in cascade and using positive feedback will have two stable states. If the inverters are replaced by multi-input circuits, the resultant binary cell can be switched from one stable state to the other by applying suitable control signals.

The simplest steered binary cell based on NAND gates is shown in Fig. 6.31. The connection between the logic gates in the cell is

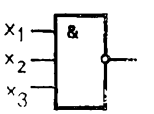


Fig. 6.30

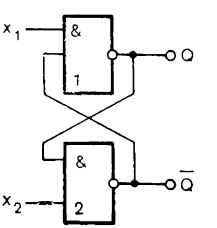


Fig. 6.31

such that

$$Q = \overline{x_1 (\overline{Qx_2})} = \bar{x}_1 + Qx_2 \tag{6.44}$$

On comparing Eq. (6.44) with Eq. (6.34), we may conclude that the binary cell in question is an asynchronous *RS* flip-flop with complementing set-reset inputs. The logic symbol for this type of flip-flop is shown in Fig. 6.25*b*.

The dynamic performance of a set-reset flip-flop cell can be evaluated in terms of two parameters, namely: $\tau_{x,\,\min}$ defined as the minimum duration of the set or reset signal required for the flip-flop to change state reliably, and τ_T defined as the total delay with which the flip-flop reaches its final output state with respect to the input signal.

In a 1-bit storage cell, both quantities are determined by the propagation delay of the signal through the two series-connected logic gates. Therefore,

$$\tau_{x1,\,\min} = \tau_{x2,\,\max} = \tau_T = 2\bar{\tau}_{pd,\,\max} \tag{6.45}$$

It is to be noted that in deriving Eq. (6.45) we practically ignore the regenerative stage of the switching process in a flip-flop consisting of logic gates. We do so because under large-signal conditions the transistors in two logic gates connected in cascade will practically never be in the active region at the same time.

Level-Type Flip-Flops with Set-Reset Inputs

Flip-flops made with NAND gates. This type of flip-flop is implemented by combining a level-type bistable cell and a combinational logic circuit which generates the functions $\bar{s}(x_1, \ldots, x_M)$ and $\bar{r}(x_1, \ldots, x_M)$ (Fig. 6.32*a*).

An asynchronous *DV* flip-flop (Fig. 6.32*b*) is synthesized in accord with Eqs. (6.35), in which *V* is replaced by *VC*. The input line for the clock pulse *C* is shown dashed in Fig. 6.32*b*.

A one-clock-input *RS* flip-flop (Fig. 6.32*c*) and a two-clock-input *D* flip-flop (Fig. 6.32*d*) have a structure answering the equations in Table 6.2. In *DV* and *RS* flip-flops, a change of state is accompanied by a consecutive switchover of three logic gates, and the total switching delay time is

$$\tau_T = 3\bar{\tau}_{pd,\,\max} \tag{6.46}$$

As for a 1-bit storage cell, the minimum duration of the trigger pulse at the *V* or *C* input is equal to $2\bar{\tau}_{pd}$ assuming that $\tau^+_{pd} = \tau^-_{pd}$. In the worst case when $\tau^+_{pd} = 2\tau^-_{pd}$ and $\tau^-_{pd} = 0$, the minimum duration of the respective pulses will approximately be

$$\tau_{V,\,\min} = \tau_{C,\,\min} = 4\bar{\tau}_{pd,\,\max} \tag{6.47}$$

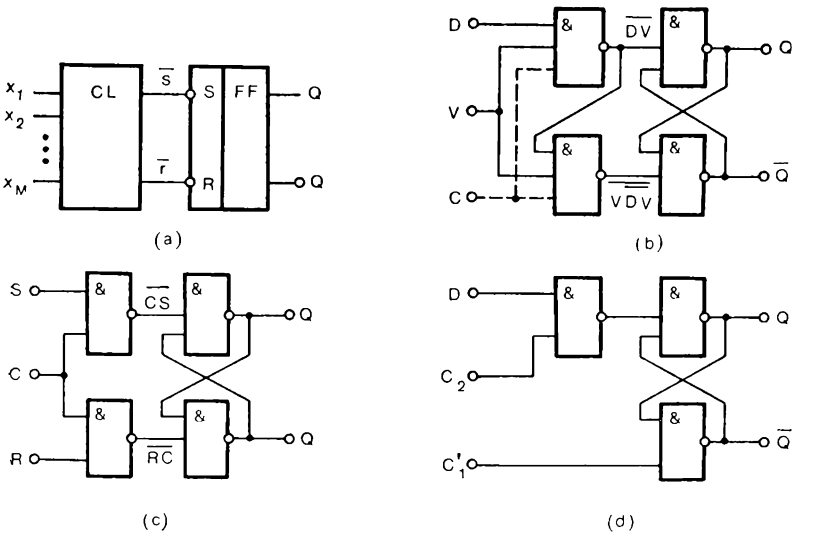


Fig. 6.32

With the clock pulse applied to the C_1' input, the dynamic behaviour of a two-clock-input D flip-flop is the same as that of a level-type binary cell [see Eq. (6.44)]. With the clock pulse applied to the C_2 input, it is similar to that of an RS flip-flop, as described by Eqs. (6.46) and (6.47).

Flip-flops made with AND-NOR gates. In these flip-flops, the functions of a level-type binary cell and a combinational logic circuit are usually combined, and the logical design of the flip-flop reduces to a superposition of the AND, OR and NOT functions of the right-hand side of the state (characteristic) equation.

A clocked normal-input RS flip-flop (Fig. 6.33*a*) can be described by re-arranging its characteristic (state) equation:

$$Q = SC + Q\overline{CR} = \overline{(\overline{C} + \overline{S})\,(\overline{Q} + CR)}$$
$$= \overline{\overline{Q}\,(\overline{C} + \overline{S}) + CR\overline{S}} = \overline{\overline{Q} + \overline{CS} + CR}$$

The equality $C\overline{RS} = CR$ stems from $RS = 0$.

A clocked complementing-input RS flip-flop (Fig. 6.33*b*) answers the following state equation:

$$Q = \overline{S'\overline{C}' + \overline{Q\overline{C}'\overline{R}'}} = \overline{(C' + S')\,(\overline{Q} + \overline{C}'\overline{R}')} = \overline{(C' + S')\,\overline{Q\,(C' + R')}}$$
$$= \overline{C'\,\overline{(QC' + QR')} + S'\,\overline{(QC' + QR')}}$$

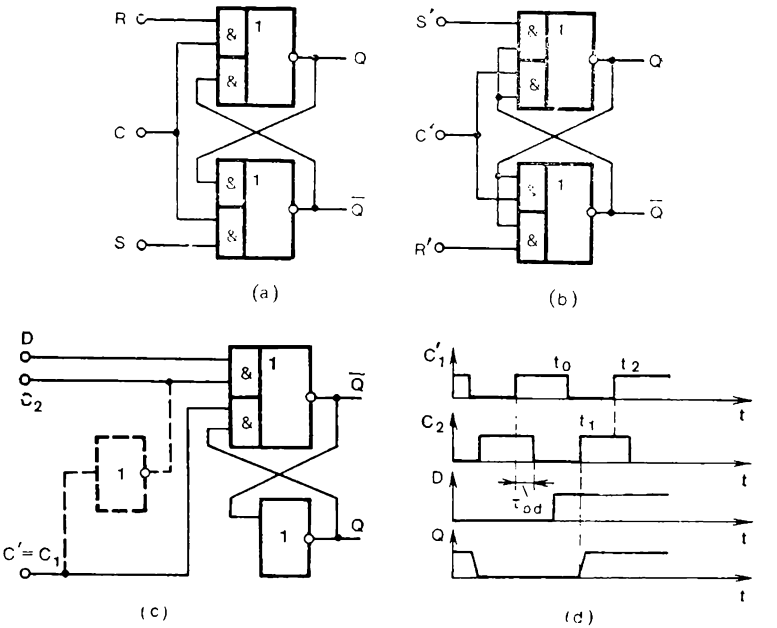


Fig. 6.33

A two-clock-input D flip-flop (Fig. 6.33*c*) answers the following state equation:

$$Q = DC_2 + QC_1' = \overline{\overline{DC_2 + QC_1}}$$

Of special interest is a clocked D flip-flop of the latch type. In our previous discussion, we have assumed that the clock pulses C_1' and C_2 occur at different instants. However, this is not the only mode of operation for a flip-flop with two clock inputs. In fact, partly overlapping clock pulses lead to a higher speed, because the reset (clear) phase partly coincides with the set phase. This principle underlies the operation of the clocked latch D flip-flop whose timing diagrams are shown in Fig. 6.33*d*. Of particular interest is the transition to the state for which $Q = D = 1$, because at $D = 0$ the input line to $C2$ is blocked and it is of no importance whether the pulses overlap or not. In the general case, for the correct operation of the flip-flop it is necessary to delay the trailing edge of C_2 relative to C_1' for a sufficiently long interval. At $\tau_{pd} = 0$, a false operation of a flip-flop made with AND-NOR gates (the transition to $Q = 0$ with $D = 1$) may only occur if the propagation delay of the positive-going edge of C_1 through the AND-NOR gate exceeds the propagation delay of the negative-going edge of C_2 by an amount such that the

false spike at the $\bar{Q}$ output is sufficiently long to cause the flip-flop to change state. This possibility is practically avoided if C_2 is the negation of C_1', in which case τ_{pd} is decided by the delay due to the inverter shown in Fig. 6.33*c* by the dashed line. With C_2 generated in this way, the flip-flop in question operates as one with a single clock input.

Flip-Flops with Internal Memory

In Sec. 6.3, we have examined a steered symmetrical flip-flop (see Fig. 6.12) which is caused to change state by a positive-going change in the input voltage. Outwardly, the characteristics of this flip-flop are fully identical with those of an asynchronous T flip-flop with a single dynamic input. A distinction of a flip-flop with steered triggering is provision of internal memory elements (memory or commutating capacitors C_{in1} and C_{in2}) which enable the previous state of the flip-flop to affect its next state.

In the general case, the next state, $Q\ (t + 1)$, of the flip-flop can be decided by either the previous state $Q\ (t)$ or the state of the binary variables at its control inputs. An important thing is that the state to which the flip-flop goes after application of the signal edge at its dynamic input depends on the state of the control inputs and the state of the flip-flop before the application of the edge. In either case, this calls for provision of additional (internal) elements in the flip-flop.

In flip-flops built from level logic gates, internal memory is provided by applying feedback to an even number of inverting logic gates. In view of the likely combinations of cell states, an asynchronous flip-flop with dynamic inputs may be treated as a multi-state sequential network which is allowed to change state at discrete instants of time determined by the delays in the generation of feedback signals. This treatment serves as the basis for the existing methods of synthesis of asynchronous sequential circuits. With them, one can implement flip-flops with any desired next-state functions [49].

A universal technique for synthesizing a flip-flop with one dynamic input is to connect in cascade two flip-flop cells having set-reset inputs. A change of state for such a flip-flop occurs in two phases, a delay phase and a transition phase. In a flip-flop with a normal dynamic input the delay phase is the time interval during which the binary input variable x_{dyn} at the dynamic input remains zero (Fig. 6.34). During this interval a bit of data is written into the first cell. The elements coupling the first and second cells are blocked, so the output signal of the flip-flop remains unchanged during the delay phase. The delay terminates and the transition commences when x_{dyn} switches from 0 to 1. This inhibits writing data into the

first cell (the set and reset functions, s and r, take on zero values), the coupling elements are unblocked, and the second cell assumes the same state as the first cell. Because no data can be written in the first cell so long as $x_{dyn} = 1$, a change of state at the control inputs cannot cause the flip-flop to change state. When x_{dyn} switches from 1 back to 0, the delay phase begins prior to a new change of state by the flip-flop.

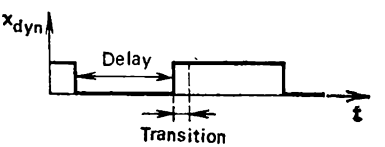


Fig. 6.34

In a flip-flop with dynamic inputs, the state of the control inputs may change in any arbitrary fashion, but the state of the first cell during the delay phase must always correspond to the last combination of the control inputs. In other words, during the delay phase the set and reset functions of the first cell must be such that

$$r + s = 1 \tag{6.48}$$

As already noted, a clocked flip-flop with internal delay may be simplified by inhibiting any change of state for the control inputs during the delay phase. The resultant flip-flop will then be analogous to a flip-flop with dynamic input. The only distinction is that the input signals will remain unchanged throughout the delay phase and the flip-flop will operate correctly even though the condition defined by Eq. (6.48) is not satisfied.

Composite flip-flops with clocked dynamic input can be characterized in terms of the following quantities:

—the minimum interval between two consecutive positive-going pulses at the dynamic input C, symbolized as $\tau^0_{C,\min}$;

—the minimum duration of the positive-going pulse at the dynamic input C, symbolized as $\tau^1_{C,\min}$;

— the total delay in the change of output state relative to the commutating edge of the signal at the dynamic input, symbolized as τ_T;

—the maximum propagation delay through the flip-flops, symbolized as $\tau_{\Sigma pd}$.

The first three quantities depend on the respective characteristics of the first and second flip-flop cells. The sum of $\tau^0_{C,\min}$ and $\tau^1_{C,\min}$ determines the maximum switching rate of the flip-flop in the case of symmetrical triggering. The last quantity is the sum of the maximum delays and defines the time interval required for the first cell to change state after the control input has changed state.

Composite flip-flops with internal delay and static clock inputs can be characterized in terms of similar quantities. The only distinction is that some of these quantities are related to the speed of the combinational logic, CL, used in the sequential circuit in question.

Also, there is a connection between the defining quantities themselves. For a flip-flop with normal clock input

$$\tau_C^0 = \tau_T + \tau_{pd\,(CL)\,\max},\quad \tau_{\Sigma pd} = \tau_T + \tau_C^1$$

For a flip-flop with a complementing clock input

$$\tau_C^1 = \tau_T + \tau_{pd(CL)\,\max},\quad \tau_{\Sigma pd} = \tau_T + \tau_C^0$$

In either case, the independent variables defining the speed of the flip-flop are τ_T and τ_C, assuming that

$$\tau_C = \begin{cases} \tau_C^1 \text{ in a flip-flop with a normal clock input} \\ \tau_C^0 \text{ in a flip-flop with a complementing clock input} \end{cases}$$

Flip-flops with cells having normal and complementing clock input. Basically, such flip-flops (Fig. 6.35) are synthesized by connecting in cascade two primitive *RS* flip-flops made with AND-NOR gates (see Fig. 6.33*a* and *b*). Now, the delay phase is the time interval during which $C = 1$. During this interval, the input gates of the first flip-flop are enabled, and its state corresponds to the values of the input signals R and S. At the same time, coupling between the first and second flip-flop is disabled, because the second flip-flop has a complementing clock input. Just as the clock signal swings in the negative direction, the input gates of the first flip-flop are disabled and the coupling elements between the first and second flip-flop are enabled. As a result, the second flip-flop assumes the same state as that of the first.

The dynamic parameters of the flip-flop described above are:

$$\tau_{C,\min}^1 = \tau_{C,\min}^0 = \tau_T = 2\bar{\tau}_{pd,\max},\quad \tau_{\Sigma pd} = 4\bar{\tau}_{pd,\max}$$

Several types of flip-flops with internal delay can be synthesized, using various primitive flip-flops with a normal clock input in the first state of the circuit in Fig. 6.35*a*.

Because primitive flip-flops built from NAND gates have a direct clock input, coupling between the stages can be controlled by inverting the clock signal. This can be done differently in different flip-flop networks.

NAND-gate *D* flip-flop with a dynamic clock input. In this type of flip-flop (Fig. 6.36*a*), the clock signal is negated by an inverter *9*. In general, it operates like an AND-NOR gate flip-flop.

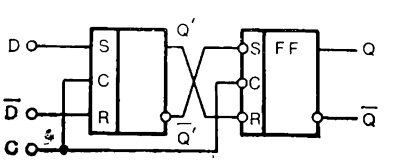


Fig. 6.35

To prevent the second-stage flip-flop from changing state during the delay phase, AND gates *5* and *6* must be disabled (blocked) before the likely change of state in the first-stage flip-flop. This is equiva-

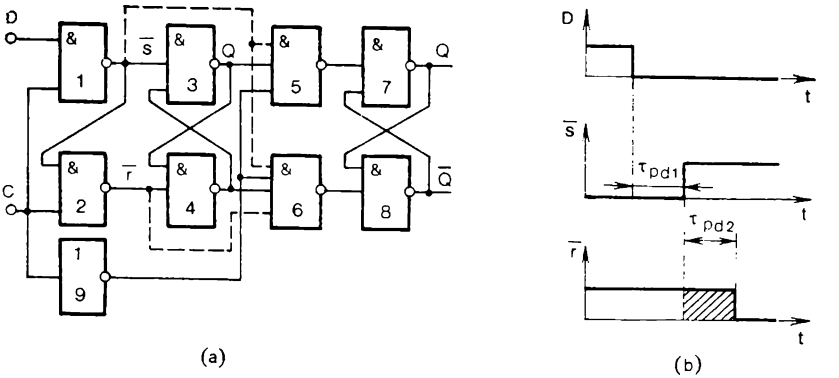


Fig. 6.36

lent to requiring that the propagation delay for the positive-going edge of the clock signal C through inverter *9* be less than the propagation delay through AND gates *1* and *3* or *2* and *4*. It should be remembered that there always is some spread in the propagation delay through logic gates and that the lower bound on this spread is usually not fixed. Obviously such a flip-flop is functionally unreliable. That is, the characteristics of the otherwise operable components may occur in combinations which might upset the normal operation of the entire flip-flop.

To avoid this, additional disabling lines (shown dashed in Fig. 6.36*a*) are provided in the flip-flop. Basically, they work as follows. If a transition of the flip-flop to the latent phase causes the first-stage flip-flop to change state, this is an indication that a 0 signal appears at the input to gate *3* or *4* before the latter has moved to the 1 state. With the additional disabling lines, gates *5* and *6* are always disabled before a 1 appears at their inputs.

We may ask if the additional disabling lines alone are sufficient for normal operation of the flip-flop, that is, can a functionally reliable flip-flop be built, omitting the inverter 9? Transient analysis would show that in a flip-flop with a dynamic clock input an inverter is essential for fundamental reasons. This is because the valid change in the input data signals during the delay phase might lead to $\bar{s} = \bar{r} = 1$ for short instants. In the D flip-flop in question, such a situation arises when D changes from 1 to 0, because gate *2* changes state with a delay relative to gate *1* (Fig. 6.36*b*). Now the first-stage flip-flop is in the $Q' = 1$ state, so it will suffice to feed a disabling signal from inverter *9* to gate *5* alone.

Flip-flops with a static clock input, in which data signals remain unchanged during the delay phase, do not need an inverter.

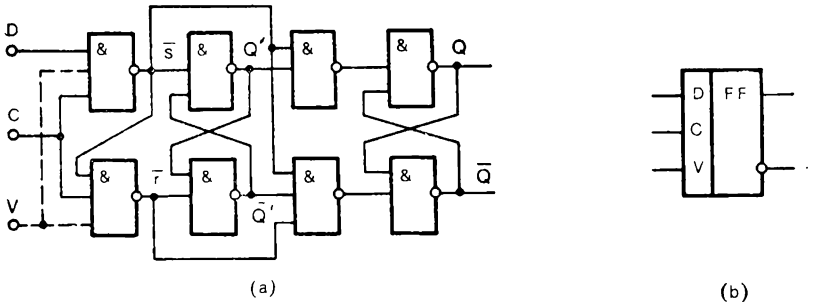


Fig. 6.37

NAND-gate *D* flip-flop with a static clock input (Fig. 6.37*a*). In this type of flip-flop, the interstage coupling lines are disabled by the $\bar{r}$ and $\bar{s}$ signals, so no inverter is needed.

NAND-gate *DV* flip-flop with a static clock input. In circuit arrangement, this type of flip-flop is analogous to a *D* flip-flop. The *V* input is formed by the disabling lines of the input gates (shown dashed in Fig. 6.37*a*). At $V = 1$, the *DV* flip-flop operates in the same manner as the original *D* flip-flop. At $V = 0$, the gates are disabled, and the flip-flop cannot change state. The logic symbol for a *DV* flip-flop is shown in Fig. 6.37*b*.

Asynchronous *JK* flip-flop. This type of flip-flop has two dynamic inputs. Because only one input can operate in each state, as asynchronous *JK* flip-flop is similar in circuit arrangement to clocked flip-flops. Consider operation of a NAND-gate *JK* flip-flop (Fig. 6.38). Suppose that initially the flip-flop is in the $Q = 0$ state. Feedback disables gate *2*, and the flip-flop can be caused to change state only by a signal at the *J* input. The delay phase is the time interval during which $J = 1$. Now the first stage assumes the $Q' = 1$ state. To disable the coupling between the flip-flop stages, an $\bar{s}$ signal is applied to gate *5*. Since the $\bar{s}$ signal remains unchanged during the delay phase, the disabling line is functionally reliable. Just as the signal at the *J* input switches from 1 to 0, gate *5* is enabled, and the flip-flop assumes the 1 state. At the same time, gate *1* is disabled whereas gate *2* is enabled to

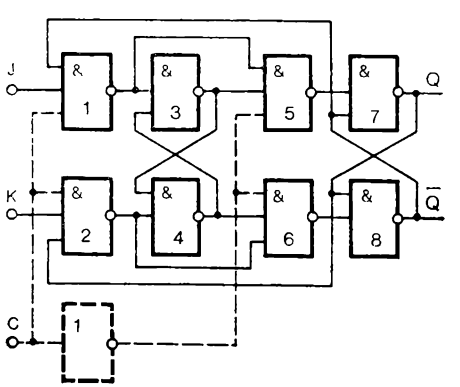


Fig. 6.38

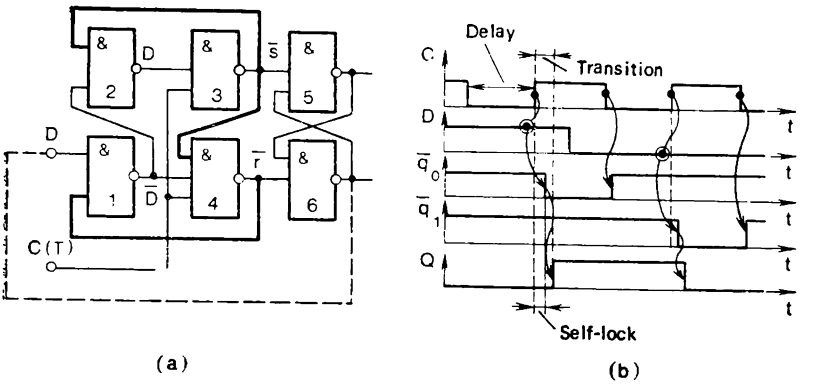


Fig. 6.39

gate the K signal. Because the circuit is symmetrical, the flip-flop changes from the 1 to the 0 state under the action of the K input in the same manner as already described.

Synchronous JK flip-flop. This type of flip-flop can be implemented by extending the above circuit to include clock lines to feed clock pulses to gates *1*, *2* and *5*, *6*. The clock pulses intended for gates *5* and *6* pass through an inverter (their path is shown by dashed lines in Fig. 6.38). The synchronous JK flip-flop operates along the same general lines as the D and DV flip-flops examined earlier.

The dynamic behaviour of NAND-gate flip-flops is determined by the dynamic characteristics of the constituent primitive flip-flops. Among other things, $\tau^0_{C,\min} = \tau^1_{C,\min} = \tau_T = 4\overline{\tau}_{pd,\max}$ and $\tau_{\Sigma pd} = 7\tau_{pd,\max}$. As is seen, NAND-gate flip-flops are not so fast as AND-NOR-gate flip-flops.

Flip-flops with a single clock input can also be synthesized in other ways.

Self-locking D flip-flop. Consider the operation of a self-locking D flip-flop built from six NAND gates (Fig. 6.39*a*). As before, each change from state to state may be visualized as consisting of two phases (Fig. 6.39*b*).

The delay phase is the time interval during which the variable at the clock input C is zero. Gates *3* and *4* are disabled, and the flip-flop formed by gates *5* and *6* retains the ability for bistable operation so that the state Q can be fixed. Because $\bar{r} = \bar{s} = 1$, the disabling lines (shown by heavy lines in Fig. 6.39*a*) do not affect the state of gates *1* and *2* which, during the delay phase, is controlled by the D input.

The transition phase commences just as the variable C goes from 0 to 1. The leading edge of the clock pulse passes through gate *4*

(if $D = 0$) or through gate *3* (if $D = 1$), and causes the output flip-flop stage to switch to the 0 or the 1 state, respectively.

An important point in the operation of a self-locking D flip-flop is to disable the line which transfers the D signal to the input of the flip-flop cell immediately after gate *3* or *4* changes state. When the flip-flop is reset to 0, gate *1* must be disabled; when it sets to the 1 state, gates *2* and *4* must be disabled. It is essential to stress that the disabling action automatically results in "memorizing" the values of $\bar{r}$ and $\bar{s}$ that exist during the positive-going edge of the clock signal—this is what may be called the self-locking or lockout action.

For the correct operation of the flip-flop, it is essential that the input variable should remain constant during short time intervals preceding and following the positive-going edge of the clock pulse. The short interval preceding the positive edge is determined by the delay in gates *1* and *2*, and the one following the positive edge by the delay in gate *3* or *4*. During the remaining time, the D variable may vary in any manner. In other words, the flip-flop in question is one with a dynamic normal clock input.

The dynamic behaviour of a self-locking D flip-flop can be described in terms of the following relations:

$$\tau^1_{C,\,\min} = 4\bar{\tau}_{pd,\,\max},\quad \tau^0_{C,\,\min} = 2\bar{\tau}_{pd,\,\max}$$

$$\tau_T = 3\bar{\tau}_{pd,\,\max},\quad \tau_{\Sigma pd} = 5\bar{\tau}_{pd,\,\max}$$

As is seen, the flip-flop in question is superior to that of Fig. 6.36 as regards both speed and the count of components needed. Among its disadvantages is that, following the edge of a clock pulse, the signal D must be kept fixed for the duration of the time delay in switching gate *3* or *4* (this time interval is required for the lockout action to take place). With several flip-flops connected in cascade this requirement might be violated if the sum of the time delays in switching the first flip-flop and the input gates (*1* and *2*) of the second flip-flop proves shorter than the lockout interval of the second flip-flop. Coupled with the spread in the dynamic characteristics of the logic gates, this drawback might make the entire sequential network functionally unreliable. This is why D flip-flops with self-locking are more frequently used as symmetrically triggered asynchronous ones. To this end, the D input is connected to the output (shown by the dashed line in Fig. 6.39*a*).

Flip-Flops with Complex Input Logic and Combined Flip-Flops

The functional capabilities of flip-flops can be extended by increasing the number of inputs.

An increase in the number of AND (OR) inputs. Consider a flip-flop made with two NAND gates (Fig. 6.40*a*). It differs from the

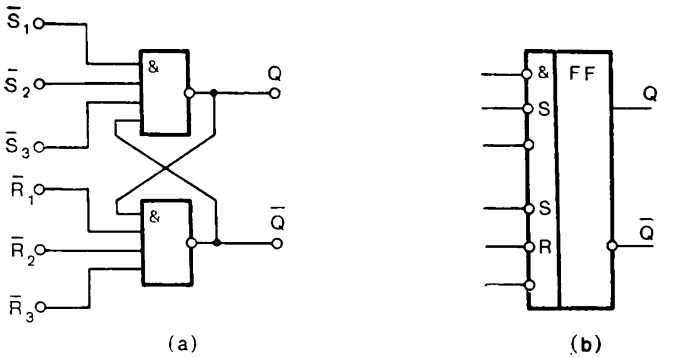


Fig. 6.40

simple flip-flop of Fig. 6.31 in that it has an increased number of inputs on the NAND gates. For the circuit to retain the capability for bistable operation, it is essential that the input variables $\overline{S}_1$, $\overline{S}_2$, $\overline{S}_3$, $\overline{R}_1$, $\overline{R}_2$, and $\overline{R}_3$ should take on value 1 all at the same time. To set the flip-flop to the 1 state it will then suffice to apply a 0 signal to at least one of the $\overline{S}_i$ inputs, and to reset it, it will suffice to apply a 0 signal to one of the $\overline{R}_i$ inputs. Then the set and reset functions may be written

$$s = S_1 + S_2 + S_3,\ r = R_1 + R_2 + R_3$$

The logic symbol for a multi-input RS flip-flop is shown in Fig. 6.40*b*.

An increased number of AND (OR) inputs may also be used on other types of flip-flops. Consider as an example the use of an increased number of D inputs on a synchronous D flip-flop. The simplest way to do this is to increase the number of inputs on the NAND gate which accepts the D signal (Fig. 6.41*a*). The resultant flip-flop may be interpreted as follows:

1. On assuming the D input to be normal, the set and reset functions take the form

$$s = CD_1D_2D_3,\ r = \overline{CD_1D_2D_3} \tag{6.49}$$

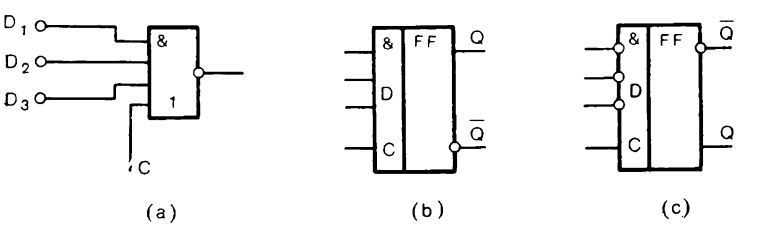


Fig. 6.41

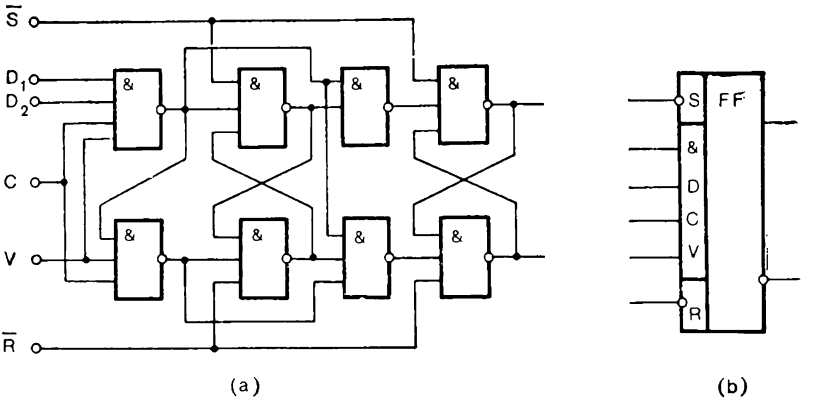


Fig. 6.42

It is seen that the flip-flop can be set to the 1 state only if a 1 signal is applied to all the D_i inputs at the same time. The logic symbol for a flip-flop with a multiple D input is shown in Fig. 6.41*b*.

2. On assuming the D input to be complementing, the previously normal (Q) output becomes a complementary ($\bar{Q}$) output, and the set and reset functions of the flip-flop are obtained by negating all the data variables, that is

$$s = C\overline{\bar{D}_1\bar{D}_2\bar{D}_3} = C\,(D_1 + D_2 + D_3)$$

$$r = C\bar{D}_1\bar{D}_2\bar{D}_3 = C\,(\overline{D_1 + D_2 + D_3})$$

The logic symbol for the flip-flop in this interpretation is shown in Fig. 6.41*c*.

In sequential circuits using D flip-flops with multiple D inputs, it is possible to transfer logic (Boolean) multiplication and addition from the combinational logic onto the flip-flop. This simplifies the network configuration and cuts down the propagation delay time through the combinational logic.

Clocked flip-flops with asynchronous set-reset inputs. In registers (see Chap. 13), it is often necessary to add asynchronous set and reset inputs to the main flip-flop inputs. This is done by increasing the number of set and reset inputs on the flip-flops. As an example, Fig. 6.42*a* shows the circuit of a single-clock-input DV flip-flop with internal delay, extended to include additional $\bar{R}$ and $\bar{S}$ inputs. The logic symbol for this flip-flop is shown in Fig. 6.42*b*. When $\bar{S} = \bar{R} = 1$, the flip-flop operates as a DV flip-flop with the inputs D_1 and D_2 and a distributed (or wired) AND connection. When a 0 signal is applied to the $\bar{S}$ (or $\bar{R}$) input, the flip-flop moves to the 1 (or 0) state, irrespective of the state at the D, V and C inputs.

Additional set and reset inputs can be provided in the other types of flip-flops in a similar way. Such flip-flops are commonly called combined. For example, there are *RST*, *JKRS*, *DRS* and *DVRS* flip-flops.

REVIEW QUESTIONS AND PROBLEMS

6.1. List methods for the synthesis of a bistable circuit.

6.2. Can a bistable circuit be built, using a CE, CB or CC transistor limiter as an amplifying element? If not, explain why.

6.3. Plot the transfer characteristic for the unsymmetrical bistable circuit of Fig. 6.3, if $E_C = 12$ V, $R_{C1} = R_{C2} = 1$ kΩ, $R_E = 200$ Ω, $R_B = 1.8$ kΩ, $R_c = 10$ kΩ, $\beta_1 = \beta_2 = 50$. The internal resistance of the signal source and the input resistance of the transistor may be neglected.

6.4. Design an unsymmetrical bistable circuit using type 1УТ401 IC opamp assuming $R_L = 5$ kΩ, $V^H = 10$ V, and $-V^L = -7$ V, if $E_{min} = -0.5$ V and $E_{max} = 1.5$ V.

6.5. Choose the rated values for the resistors used in the symmetrical saturated bistable circuit of Fig. 6.6, if $P = 100$ mW, $E_C = 12$ V, $E_B = 1$ V, $e_{E0} = 0.3$ V, $\beta_{min} = 40$, and $I_{CEO, max} = 100$ μA.

6.6. Determine the minimum value of R_L connected in parallel with R_C, at which the bistable circuit of Problem 6.5 retains its capability for correct operation.

6.7. Find the minimum values of E_C for the bistable circuit of Fig. 6.7, if $R_C = 1$ kΩ, $R_{L,min} = 2$ kΩ, and $E_{c1} = 10$ V.

6.8. List the phases in the switching process for a symmetrical saturated transistor bistable circuit with asymmetrical base triggering.

6.9. Construct an approximate plot for the minimum amplitude of a trigger (input) pulse as a function of its duration.

6.10. Explain how the saturation factor of the ON transistor in a symmetrical bistable circuit affects the required duration of the trigger pulse.

6.11. List the inputs of elementary sequential machines and explain their functions.

6.12. Outline the differences between asynchronous and synchronous flip-flops.

6.13. How can the next-state (transition) function of an asynchronous flip-flop be specified? Explain the difference.

6.14. The next-state function of an asynchronous flip-flop is $dQ = Qx_1x_2\,dx_2 + \bar{Q}x_1x_3\,dx_3$. Name the types of inputs x_1, x_2 and x_3. Describe verbally the operation of the flip-flop and propose a logic symbol for it.

6.15. Write the next-state function of Problem 6.14 in normal notation.

6.16. Give the types and logic symbols for the flip-flops specified by the equations:

(a) $dQ = x_1(x_2 \oplus Q)\bar{x}_3\,dx_3,\ x_3\,dx_1 = 0$

(b) $dQ = \bar{Q}x_1\,dx_1 + Qx_2\,dx_2,\ \bar{x}_1\bar{x}_2 = 0$

(c) $dQ = x\,dx$

(d) $dQ = (x_1 \oplus Q)\,x_2\,dx_2$

(e) $dQ = (Qx_1 + \bar{Q}x_2)\,x_3\,dx_3,\ x_3\,dx_1 = x_3\,dx_2 = 0$

6.17. Compare in terms of speed sequential networks based on clocked flip-flops with a dynamic clock input and those based on internal-delay flip-flops with a static clock input.

6.18. Draw the circuit of an asynchronous complementing-input *RS* flip-flop implemented with TTL gates incorporating a totem-pole output drive. How should the circuit be re-arranged to obtain normal inputs?

6.19. Draw logic diagrams for asynchronous normal and complementing input *T* flip-flops based on AND-NOR gates, and circuit diagrams for similar flip-flops using basic TTL gates.

6.20. Draw a logic diagram for an internal-delay *D* flip-flop based on AND-NOR gates and also a circuit diagram for a similar flip-flop based on MOS transistors and inverters with a quasi-linear load. Find out if the clock input is dynamic.

7 MULTIVIBRATORS

7.1. GENERAL

If in a symmetrical pulse-shaping bistable circuit we make one or both coupling connections purely capacitive, one or both stable states will become unstable, and the time during which the circuit will be able to reside in them will depend on the charging and discharging of the coupling capacitors. These unstable states are characterized by slow variations in currents and voltages (as compared with regeneration processes), so they are often called *temporary* or *quasi-stable*. Pulse generators in which interstage coupling is provided by resistive-capacitive elements and which can reside in one or two quasi-stable states are called *multivibrators**.

Multivibrators can operate in the monostable (one-shot), astable (free-running), or synchronized mode.

A *monostable* (*one-shot*) *multivibrator* has, as its name implies, one stable state in which it can reside for an indefinitely long time, until a triggering pulse arrives to switch it to the other state. The other state is only temporarily stable and can last for a predetermined interval τ, after which the circuit returns to its original state.

An *astable* (*free-running*) *multivibrator* has two momentarily stable states between which it continuously alternates, remaining in each for a period controlled by the circuit parameters. The period of switching, T, between the two states is the sum of τ_{01} and τ_{02}, that is the times during which the circuit remains in either temporary state.

A *synchronized multivibrator* switches from the temporary to the permanently stable (or another unstable) state at an instant tied to a particular time position (phase) of a periodic synchronizing signal. Operation of a synchronized multivibrator is taken up in more detail in Chap. 10.

Multivibrators can be built with transistors, tubes [4], negative-resistance devices (see Chap. 12), and integrated circuits.

---

* In the US and UK literature of the subject, such circuits are usually referred to as *bistable multivibrators*. In the Soviet literature, however, they are not included with multivibrators at all.—Translator's note.

7.2. DISCRETE-COMPONENT MULTIVIBRATORS

Collector-Coupled Monostable Multivibrator

The circuit of and the waveform generated by a collector-coupled monostable multivibrator are shown in Fig. 7.1. One transistor, $T2$, is normally saturated, and the other transistor, $T1$, is at cutoff. In a temporary state, $T2$ is at cutoff and $T1$ is at saturation. Positive trigger pulses are usually applied to the collector of $T1$ or the base of $T2$. The events initiated by triggering and at the first transition (time t_0) are analogous to those taking place in an asymmetrically triggered bistable multivibrator.

The duration of the temporary state is controlled by the voltage existing on the base of $T2$ as the capacitor C discharges. This time-setting voltage decays exponentially (Fig. 7.1*b*). The temporary state terminates at time t_1, when the time-setting voltage reaches the turn-on value for $T2$ (in Fig. 7.1*b*, this voltage is taken to be approximately equal to zero). The collector current of $T2$ moves $T1$ out of saturation, and a regenerative process takes place in the circuit which terminates by causing $T1$ to turn off. As in a bistable multivibrator, this is accompanied by a relatively fast (nearly stepwise) positive change in the collector voltage of $T2$.

At $t > t_1$, the initial voltage on the capacitor C is restored, which is accompanied by an exponential negative change in the collector voltage of $T1$. Therefore, the duration of the negative-going edge of v_{C1} may be taken as the recovery time, that is the time required for the circuit to go back to its original state: $\tau_{e1} \approx \tau_{re}$.

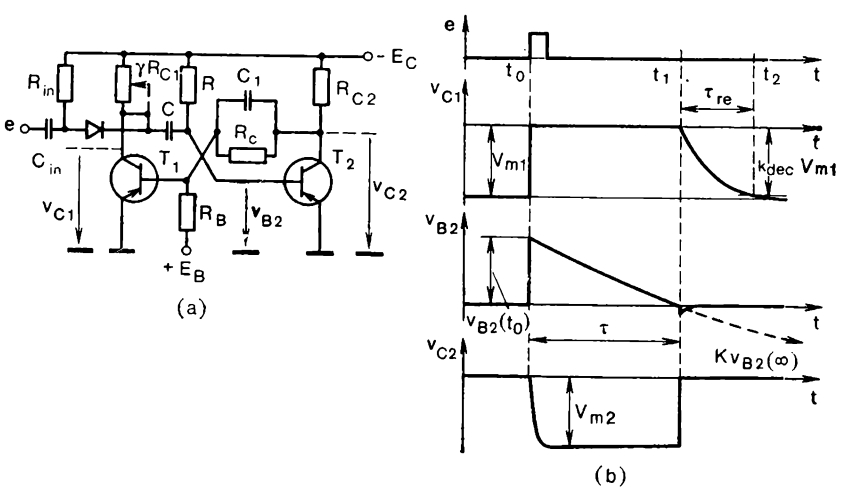


Fig. 7.1

The amplitudes of the voltage pulses on the collectors of $T1$ and $T2$ are given by

$$V_{m1} = E_C - I_{CBO1}R_{C1} - V_{CE(sat)} \tag{7.1}$$

$$V_{m2} = R_c (E_C - I_{CBO2}R_{C2})/(R_c + R_{C2}) \tag{7.2}$$

Operability conditions. The circuit components T_1, R_{C1}, R_c, R_B, E_B, and C_1 make up an inverter for which the operability conditions are defined by Eqs. (5.19) and (5.20). It should be noted that in a monostable multivibrator R_{C1} and R_{C2} may have different values. Also, the collector saturation current of $T1$ exceeds the current through R_{C1} by an amount equal to the capacitive current i_C the maximum value of which is close to $2E_C/R$. This factor can be accounted for by substituting $\beta_{eq.min} = \beta_{min}R_{C1}/R_{C2}\,(1+2R_{C1}/R)$ in Eq. (5.20).

For $T2$ to be at saturation during the stable (initial) state the value of the base resistor is chosen such that

$$R \leqslant R_{max} = \beta_{min}R_{C2} \tag{7.3}$$

For $T2$ to turn off at the beginning of the temporary state, it is essential that the initial step change in the base voltage of $T2$ exceed the turn-off voltage of that transistor. With a sufficiently large value of the timing capacitor, this step change may be deemed equal to the change of collector voltage on $T1$: $v_{B2}(t_0) \approx E_C$. Thus, $T2$ will reliably turn off, provided $E_C > e_{EO}$ which is always the case in practice.

Duration of a temporary state (the interval from t_0 to t_1). As a rule, this interval markedly exceeds the rise and fall times, so it may be taken as the duration τ of a multivibrator pulse. As already noted, it is controlled by the exponentially decaying timing voltage

$$v_{B2}(t) = v_{B2}(\infty) - [v_{B2}(\infty) - v_{B2}(t_0)] \exp[-(t-t_0)/\tau_1] \tag{7.4}$$

where $\tau_1 \approx CR$, $v_{B2}(t_0) = V_{m1} - e_{EO} = E_C - I_{CBO}R_{C1} - e_{EO} - V_{CE(sat)}$, $v_{B2}(\infty) = -E_C - I_{CBO}R$, $V_{CE(sat)}$ and e_{EO} are the parameters of the simplified piecewise-linear transistor model (see Sec. 3.2) defining the respective collector voltage of $T1$ and base voltage of $T2$ in the saturated region.

On setting $t = t_1$ in Eq. (7.4) and substituting $v_{B2}(t_1) = -e_{EO}$, we obtain an equation whose solution for the interval $\tau = t_1 - t_0$ has the form

$$\tau = k_\tau \tau_1 \tag{7.5}$$

where

$$k_\tau = \ln\left[2 + \frac{e_{EO} - I_{CBO}(R + R_{C1}) - V_{CE(sat)}}{E_C + I_{CBO}R - e_{EO}}\right] \tag{7.6}$$

Discarding the first-order infinitesimals in Eq. (7.6) leaves $\tau \approx \tau_1 \ln 2 \approx 0.7\tau_1$.

It is seen from the above relations that the value of τ is mainly affected by the coupling capacitor C and the base resistor R which together serve as the time-setting (timing) components of the multivibrator. The correction terms e_{EO}, I_{CBO} and $V_{CE(sat)}$ in Eq. (7.6) are considered when evaluating the effect of the transistor parameters and the supply voltage on τ. Equation (7.6) shows that the variations in τ because of variations in I_{CBO} and e_{EO} with temperature decrease with increasing E_C.

The value of τ can be controlled at will by varying any one of the following: τ_1, $v_{B2}(t_0)$, or $v_{B2}(\infty)$ [see Eq. (7.4)].

With C held constant, τ_1 can continuously be controlled if R is a variable resistor. Noting that $\tau \sim R$, the maximum to minimum duration (span) ratio is

$$k_s = \tau_{max}/\tau_{min} - 1 = R_{max}/R_{min} - 1 \tag{7.7}$$

Since R_{max} is defined by Eq. (7.3), for k_s to increase we must reduce R_{min}. However, a decrease in R would entail an increase in the current of $T1$ in a temporary state and also in the turn-off delay of $T2$, so that the sensitivity of the circuit towards trigger pulses would be impaired. Therefore, it is usual to choose $k_s \leqslant 1$ for this method of continuous pulse duration control.

The value of $v_{B2}(t_0)$ is usually controlled by tapping the timing capacitor C down on the collector load resistor R_{C1} (the connection is shown by the dashed line in Fig. 7.1*a*). Then the step change in the base voltage of $T2$ will be determined by the setting of the potentiometer which controls the division ratio of the voltage divider, $\gamma = R'_{C1}/R'_{C1}$, where R'_{C1} is the upper part of the collector load resistor:

$$v_{B2}(t_0) = -e_{EO} + V_{m1}\gamma \tag{7.8}$$

Substituting Eq. (7.8) in Eq. (7.4) gives

$$k_\tau = \ln\left[1 + \gamma \frac{E_C - I_{CBO}R - V_{CE(sat)}}{E_C - e_{EO} + I_{CBO}R_{C1}}\right] \approx \ln(1+\gamma) \tag{7.9}$$

The span ratio for this method of pulse duration control is limited by the minimum value of γ for which the feedback loop gain of the multivibrator remains greater than unity. Practically, k_s may be 10 and higher.

The value of $v_{B2}(\infty)$ can likewise be varied by connecting the upper end of the timing resistor R to the control voltage source $-E_1$. Then, $v_{B2}(\infty) = -E_1 - I_{CBO}R$. An advantage of this method of continuous pulse duration control is that the multivibrator can be controlled remotely, using a suitable electric signal.

Recovery (resolving) time (the interval from t_1 to t_2). The duration of this interval is determined by the exponential build-up of the charge on the timing capacitor due to the current flowing through R_{C1}. Therefore, we may write

$$\tau_{re} = CR_{C1}k_{re} \tag{7.10}$$

where k_{re} is the recovery factor in turn controlled by the decay level k_{dec} of the exponential process:

$$k_{re} = -\ln(1 - k_{dec}), \quad k_{dec} = |v_{C2}(t_2)/V_m| \tag{7.11}$$

Taking $k_{dec} = 0.95$ for the circuit of Fig. 7.1a, we have $k_{re} = 3$.

The sum $\tau + \tau_{re}$ determines the minimum repetition period, T_{min} for trigger pulses. Should trigger pulses be applied before the recovery time elapses ($\tau < T < T_{min}$), the residual voltage on the timing capacitor and, as a consequence, the duration of the multivibrator pulses would depend on the repetition period of trigger pulses, which is usually to be avoided. If this were allowed to happen, the circuit would move to one of its temporary states. In one of them, the pulse duration would then be shorter than under normal conditions. In the other temporary state, the circuit would alternately generate longer and shorter pulses.

As a way of improving the trailing edge of the collector voltage pulse v_{C1} (without having to cut down the recovery time), it is usual to couple the timing capacitor C to the collector load resistor R_{C1} via an additional resistor and a turn-off diode D (Fig. 7.2). In the temporary state, the diode is conducting and permits the timing capacitor to discharge. When the multivibrator goes back to its original state, the diode turns off, and the capacitor charges via R'_{C1} whose value is ordinarily chosen to be equal to R_{C1}. As a result, the fall time of v_{C1} is cut down substantially and approaches the pulse rise time.

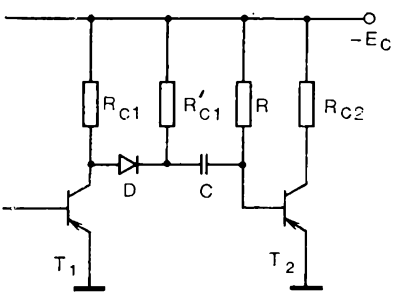


Fig. 7.2

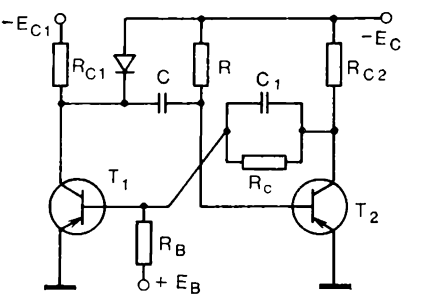


Fig. 7.3

An improvement in the speed of the circuit owing to a cut-down in the recovery time can be achieved by including a cutoff diode or an emitter follower in the charging circuit of the timing capacitor.

In a multivibrator with a cutoff diode (Fig. 7.3), the capacitor charges up while the diode is at cutoff, from the $-E_{C1}$ supply chosen such that $|E_{C1}| > |E_C|$. The capacitor ceases charging when v_{C1} has dropped to $-E_C$ and the diode turns on, thereby clipping the final portion of the exponential by varying charge. Because of this, the value of k_{re} in Eq. (7.10) decreases to

$$k_{re} = -\ln (1 - E_C/E_{C1}) \tag{7.12}$$

When an emitter follower is placed between $T1$ collector and $T2$ base, the return of the multivibrator to its original state causes the timing capacitor to charge up via the low output resistance of the emitter follower rather than the first-stage load.

Astable (Free-Running) Multivibrators

The basic circuit of a collector-coupled astable (free-running) multivibrator is shown in Fig. 7.4*a*. As in a monostable multivibrator, the ON transistor must be in saturation so as to stabilize the pulse amplitude which governs the period of oscillation. One problem with the collector-coupled astable circuit is that it may not always start oscillating when the supply voltage is turned on. Because of the circuit symmetry, it may so happen that both transistors switch on and remain on, and the circuit is in a stable state; a trigger pulse will then be necessary to drive the circuit into the relaxation mode.

During the first half-cycle (Fig. 7.4*b*). $T2$ is in saturation, and $C2$ discharges through $R2$. At the same time, $T1$ is at cutoff due to the

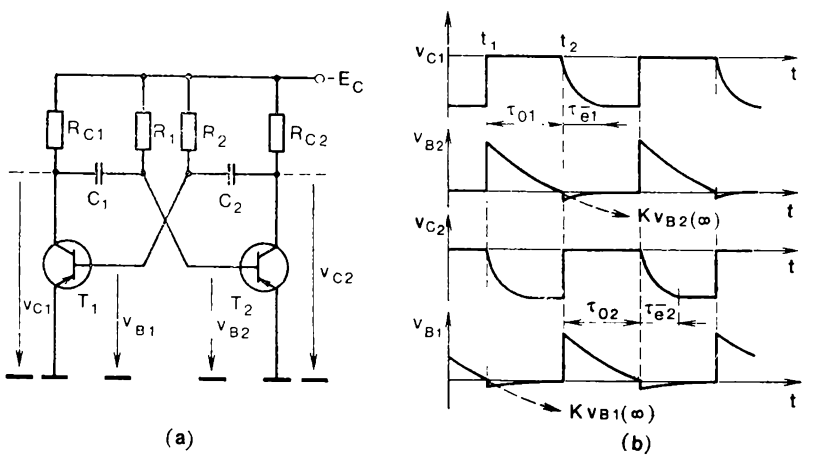


Fig. 7.4

positive timing voltage v_{B1} taken from the left-hand plate of $C2$. This temporary state terminates (at time t_1) when the timing voltage rises to the turn-on value for $T1$. The collector current of $T1$ pulls $T2$ from saturation because, so long as the emitter junction of $T2$ is forward-biased, the bulk of collector current flows through $C1$ and the emitter junction of $T2$. As a result, the feedback loop gain exceeds unity, and a cumulative switchover takes place in the circuit, which terminates in $T2$ turning off. This is followed by a rapid positive-going step change in the collector voltage of $T1$ because its collector current rises to its maximum value. The rising collector voltage of $T1$ coupled to $T2$ base determines the initial value of the timing voltage v_{B2} which holds $T2$ at cutoff during the other temporary state. The negative-going change of collector voltage on $T2$ as it turns off depends on the charging of $C2$ via R_{C2} and the emitter junction of the saturated transistor $T1$. The second half-cycle terminates when the timing voltage has reduced to the turn-on value for $T2$. At that instant, a second switchover takes place, and the multivibrator goes back to the original temporary state considered at the beginning.

Operability conditions. For the ON transistor to be in saturation, it is required that

$$I_{B1(2)} \geqslant i_{C1(2)max}/\beta_{1(2)min} \tag{7.13}$$

As in a monostable multivibrator, the collector current of the saturated transistor is not limited by the ratio E_C/R_C, but has a second component which flows through the CR network in the base circuit of the OFF stage. Neglecting the relatively small components containing the initial collector current, Eq. (7.13) may be re-written as

$$R_{1(2)} \leqslant \beta_{min} R'_{C1(2)}$$

where $R'_{G1(2)} = R_{C1(2)} \parallel 0.5R_{2(1)}$.

The condition for the other stage to be turned off by the initial (positive) change of collector voltage in the first stage that has already moved into saturation is satisfied practically always because the cutoff voltage e_{EO} is negligibly small.

Period of oscillation. Neglecting the pulse rise and fall times, the period of oscillation T may be written as the sum of two half-cycles:

$$T = \tau_{01} + \tau_{02}$$

each of which is desermined by hhe timing voltage on the base of the OFF transistor. By analogy with Eq. (7.5), the period T may be written

$$T = k_{\tau 1}\tau_1 + k_{\tau 2}\tau_2 \tag{7.14a}$$

where $\tau_{1(2)} = C_{1(2)}R_{1(2)}$, and $k_{\tau 1(2)}$ can be found by Eq. (7.6) on substituting $R = R_{1(2)}$. Assuming that the two arms of the multi-

vibrator are identical, we get, to a first approximation

$$T \approx 1.4CR \tag{7.14b}$$

Equations (7.14) hold if the pulse has time to reach a steady-state amplitude during each half-cycle, which is equivalent to satisfying the condition

$$\tau_{e1(2)} \approx 3C_{1(2)}R_{C1(2)} \leqslant \tau_{01(2)} \tag{7.15}$$

Equation (7.15) sets the limit for the maximum pulse period-to-pulse duration ratio (the reciprocal pulse duty) in case the symmetry of the multivibrator is upset.

7.3. MODIFIED (INTEGRATED CIRCUIT) MULTIVIBRATORS

In circuit configuration, integrated-circuit (IC) multivibrators are not unlike multivibrators using discrete circuit components. Examples of IC multivibrators are the Soviet-made Series 119 type К1ГФ192 and К1ГФ193 and Series 218 type К2ГФ181 and К2ГФ182. Among their features are:

—provision of turn-off diodes in the collector leads to cut down the rise and fall times of collector pulses;

—"soft" excitation owing to nonlinear negative feedback which prevents transistor saturation;

—provision of diode limiters in the base leads to prevent the breakdown of the emitter junction in integrated drift transistors.

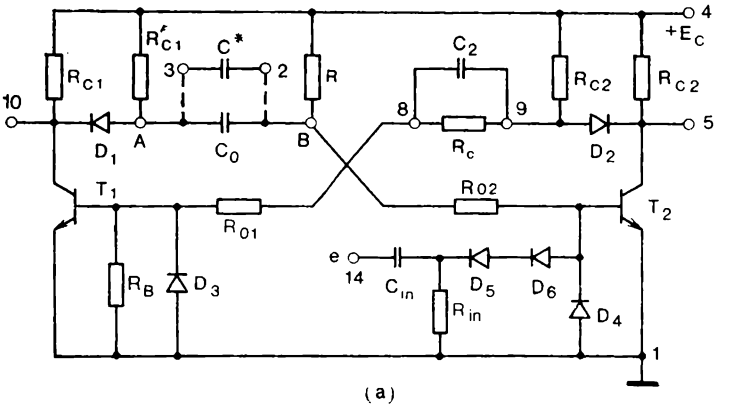
IC Monostable Multivibrators

We shall analyze the operation of an IC monostable multivibrator with reference to a Series 218 type К2ГФ182 unit (Fig. 7.5*a*). To ensure monostable operation, a discrete resistor R_C^* is placed between pins *8* and *9*, whereas a discrete capacitor C^* (connected between pins *2* and *3*) in the time-setting circuit determines the required duration τ for the temporary state. $D1$ and $D2$ are turn-off diodes, whereas $D3$ and $D4$ are parallel limiters.

In the stable state, $T2$ is saturated and $T1$ is at cutoff. C_0 and C^* are charged up to

$$V_{C_0} = V_{AB} = (E_C - e_{EO})R/(R + R_{02})$$

The negative-going trigger pulse applied at time t_0 (Fig. 7.5*b*) switches the multivibrator to the temporary state; $T1$ moves into the saturated region and $T2$ turns off. The voltage at points A and B suddenly drops by $V_m = E_C - e_{01} - V_{CE(sat)}$, limiter diode $D4$ jumps into conduction, and a negative voltage $v_{B_2}(0) = -e_{04}$ appears at $T2$ base. In the temporary state, the relaxation process may be visualized as proceeding in two stages. During the first stage



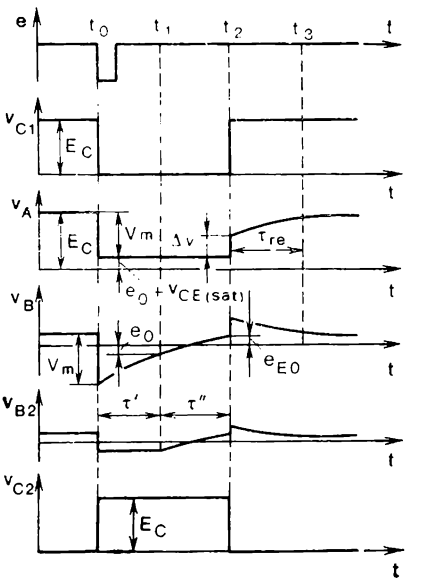


Fig. 7.5

($\tau' = t_1 - t_0$), $D4$ is conducting, the voltage at point B rises exponentially to $v_B(t_1) = -e_{04}$, such that $D4$ turns off, and the second stage of the relaxation process ($\tau'' = t_2 - t_1$) commences. During this stage, the capacitive current decreases and the exponential build-up of the potential at point B slows down. The second stage terminates at time t_2 when the exponential voltage at point B reaches

the turn-on threshold value, $v_B(t_2) = v_{B2}(t_2) = e_{EO}$, for $T2$. This causes the multivibrator to switch back to its original stable state where $T2$ is saturated and $T1$ is at cutoff. At $t > t_2$, the circuit begins to recover its initial state—the timing capacitor $C = C_0 + C^*$ begins to charge.

The value of the discrete resistor R_c^* is chosen so as to ensure the correct operation of $T1$. For $T1$ to be at cutoff in the stable state, it is required that

$$v_{B1} = \frac{R_B(V_{CE(sat)} + e_{EO2})}{R_c^* + R_{O1} + R_B} \leqslant e_{EO1}$$

The minimum value of R_c^* is thus chosen subject to the above inequality. For $T1$ to be in saturation in the temporary state, it is required that

$$i_{B1} \geqslant I_{C(sat)max}/\beta_{1,\,min}$$

This inequality defines the maximum value for R_c^*. The optimal value of R_c^* is usually given in data sheets for the respective integrated circuits.

Duration of the temporary state. The duration of this interval is determined by the timing voltage at point B. The manner in which the timing voltage varies can be ascertained from the equivalent circuit of Fig. 7.6*a*. The state of the switch (Sw) is decided by the limiting diode $D4$. During the first stage in the relaxation process, the switch is ON, whereas during the second state it is OFF. The equivalent circuit has been developed, using simplified piecewise-linear models for the diodes and transistors. Because in integrated circuits the function of diodes is performed by the emitter junctions of the IC transistors, we may assume that $e_{01} = e_{04} = e_{EO2} = e_0 \approx 0.6$ V. Then the duration of the temporary state will be

$$\tau = \tau' + \tau'' = \tau_1 k_\tau' = \tau_2 k_\tau''$$

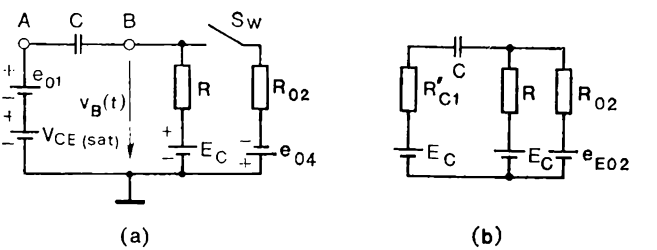


Fig. 7.6

where

$$\tau_1 = (C_0 + C^*)\, RR_{02}/(R + R_{02})$$

$$\tau_2 = (C_0 + C^*)\, R$$

$$k'_\tau = \ln \frac{v_{B1}(\infty) - v_B(t_0)}{v_{B1}(\infty) - v_B(t_1)}$$

$$k''_\tau = \ln \frac{v_{B2}(\infty) - v_B(t_1)}{v_{B2}(\infty) - v_B(t_2)}$$

and also

$$v_B(t_0) = -\frac{E_C R}{R + R_{02}} \left[1 - \frac{e_0 + (e_0 + V_{CE(sat)})\,(1 + R_{02}/R)}{E_C} \right]$$

$$v_B(t_1) = -e_0$$

$$v_B(t_2) = e_0$$

$$v_{B2}(\infty) = E_C$$

$$v_{B1}(\infty) = (E_C R_{02} - e_0 R)/(R + R_{02})$$

Subject to

$$\frac{e_0 + V_{CE(sat)} + 2e_0/(1 + R_{02}/R)}{E_C + e_0} \ll 1 \text{ and } (e_0/E_C) \ll 1$$

the expressions for k'_τ and k''_τ may be simplified

$$k'_\tau \approx \ln (1 + R/R_{02}) - \frac{e_0 + V_{CE(sat)} + 2e_0/(1 + R_{02}/R)}{E_C}$$

$$k''_\tau \approx 2e_0/E_C$$

In the case of a type К2ГФ182 integrated circuit for which [53]: $R_{02} = 1$ kΩ, $R = 15$ kΩ, $E_C = 6.3$ V, $e_0 = 0.6$ V, and $V_{CE(sat)} = 0.2$ V, we get: $k'_\tau = 2.56$ and $k''_\tau = 0.19$. Hence,

$$\tau = (C_0 + C^*)\, R\, [k'_\tau/(1 + R/R_{02}) + k''_\tau]$$

$$= 0.35\,(C_0 + C^*)\, R \qquad (7.16)$$

Equation (7.16) can be used to find the value for the discrete capacitor C^*. The value of C_0 is found according to the data sheet value of τ for $C^* = 0$, namely: $C_0 = \tau_{data}/0.35R$. For К2ГФ182 integrated circuits, $\tau_{data} = 0.9$ to 1.5 μs. Hence, $C_0 = 180$ to 300 pF.

The recovery (resolving) time τ_{re} of the circuit depends on the exponential charging of the timing capacitor C. Referring to the equivalent circuit for the charging circuit shown in Fig. 7.6*b*, the equivalent time constant for the charge recovery by the capacitor is seen to be

$$\tau_{eq} = C\,[R'_{C1} + RR_{02}/(R + R_{02})]$$

On choosing the exponential decay factor equal to $k_{dec} = 0.95$, the recovery (resolving) time is found to be $\tau_{re} = 3\tau_{ed}$.

IC Free-Running Multivibrators

As an example, consider the operation of a К1ГФ192 integrated circuit unit (Fig. 7.7*a*). Its waveforms at characteristic points appear in Fig. 7.7*b*. The IC in question is a combination of two self-contained inverters whose outputs are cross-coupled via external capacitors $C1$ and $C2$ to the opposite inputs. The turn-off diodes $D1$ and $D2$ are provided to improve the waveform of collector voltage pulses. The emitter junctions are safeguarded against a breakdown, because $E_C = 3$ V is less than V_{bd}.

The use of unsaturated inverters with nonlinear negative feedback provides for a soft excitation of the circuit. To this end, the resistors in the base and collector leads of the inverters are chosen such that, when the inverters operate on their own, the diodes $D3$ and $D4$ in the feedback circuits are forward-biased when the respective transistors $T1$ and $T2$ are ON.

The period of oscillation, T, of the multivibrator is the sum of the time intervals τ_{01} and τ_{02} during which $T1$ and $T2$ are in turn at cutoff. The value of τ_{01} is found, assuming that while $T1$ is at cutoff, the timing voltage on the capacitor $C1$ varies exponentially from the initial value $V_{C1,0} = E_C - e_{E01}$ to its steady-state value $V_{C1,\infty} = E_C R_{B1}/(R_1 + R_{01} + R_{B1})$. Then,

$$\tau_{01} = k_{\tau 1}\tau_1 \tag{7.17}$$

where

$$\tau_1 = C_1 R_{1,eq}$$

$$k_{\tau 1} = \ln 2 + \ln\left(1 + \frac{R_1 + R_{01}}{2R_{B2}}\right) + \ln\left(1 - \frac{e_{02} - e_{E01} + V_{CE(sat)}}{(1+\gamma)E_C}\right) - \ln\left(1 - e_{E01}/\gamma E_C\right)$$

where

$$\gamma = R_{B1}/(R_1 + R_{01} + R_{B1})$$

$$R_{1eq} = (R_1 + R_{01}) \parallel R_{B1}$$

Because the IC in question is such that

$$\frac{e_{02} - e_{E01} + V_{CE(sat)2}}{(1+\gamma)E_C} \ll 1$$

and

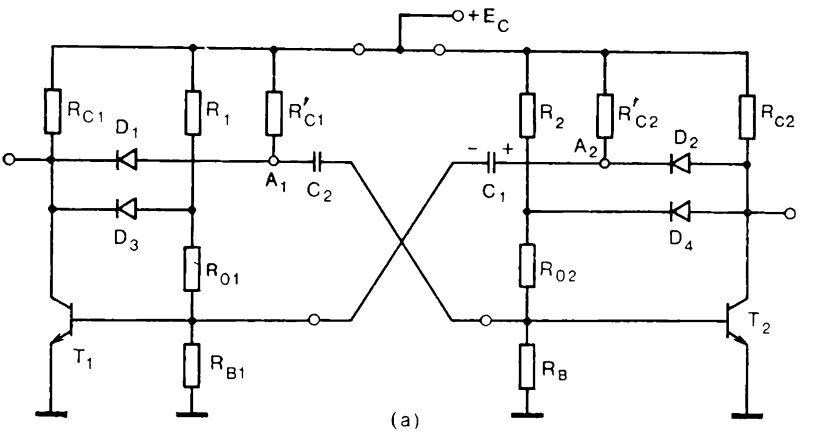
$$e_{E01}/\gamma E_C \ll 1$$

the expression for $k_{\tau 1}$ may be simplified

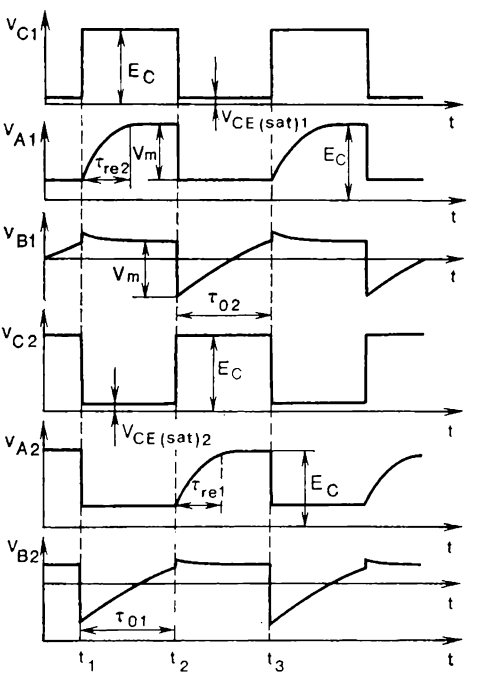
$$k_{\tau 1} \approx 0.7 + \ln\left(1 + \frac{R_1 + R_{01}}{2R_{B1}}\right) - \frac{e_{02} - e_{E01} + V_{CE(sat)}}{(1+\gamma)E_C} - e_{E01}/\gamma E_C$$

$$\approx 0.7 + \ln\left(1 + \frac{R_1 + R_{01}}{2R_{B1}}\right) \tag{7.18}$$

The value of τ_{02} can be found from Eqs. (7.17) and (7.18) on replacing the circuit parameters of the first inverter with those of the second.



(a)



(b)

Fig. 7.7

Owing to the symmetry of the IC, it may be deemed that $R_{1,eq} = R_{2,eq} = R_{eq}$ and $k_{\tau 1} = k_{\tau 2} = k_T$.

The period of oscillation of the multivibrator in question is

$$T = \tau_{01} + \tau_{02} = k_T (C_1 + C_2) R_{eq} \tag{7.19}$$

The above relation holds true on the condition that the voltages on C_2 and C_1 are fully restored over the intervals equal to τ_{01} and τ_{02}, respectively, which can take place if

$$\tau_{re,2} \approx k_{re} C_2 R'_{C1} \leqslant \tau_{01} = k_T C_1 R_{eq}$$

and

$$\tau_{re,1} \approx k_{re} C_1 R'_{C2} \leqslant \tau_{02} = k_T C_2 R_{eq}$$

For a К1ГФ191 integrated circuit which has $k_T = 1$ to 1.2, $k_{re} = 3$, and $R_{eq}/R'_{C1(2)} \approx 3$ to 3.5, Eq. (7.19) holds true at $Q_C \leqslant 1.8$ to 2.2. Incomplete recovery of voltages on C_2 and C_1 entails a reduction in $V_{C1,0}$ and $V_{C2,0}$ and in the values of τ_{01} and τ_{02}. In such a case, the period of oscillation is found as explained in [53].

7.4. IC LOGIC GATE MULTIVIBRATORS

In cases where a multivibrator is to operate as part of a digital network assemblied from IC logic gates and no stringent requirements are imposed on the time stability of the generated pulses, it is advantageous to build its switches from the same components as the digital part of the complete network. With this approach, standard components can be used throughout the equipment and a natural compatibility is achieved between the amplitude characteristics of the multivibrator pulses and the signals in the digital part of the network.

Consider the principles underlying the analysis and synthesis of the basic types of multivibrators with reference to positive-logic TTL with a totem-pole output driver (see Chap. 5). These principles can readily be extended to other types of IC logic gates.

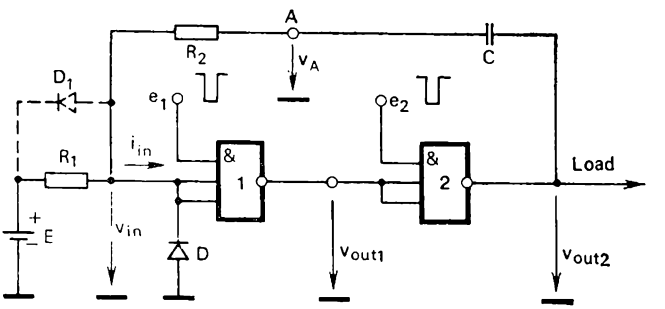


Fig. 7.8

A single-shot multivibrator can readily be derived from an asynchronous RS flip-flop (see Fig. 6.31) on replacing one of the direct (resistive) coupling circuits by a capacitive one as shown in Fig. 7.8. The status of the logic gate in the initial (permanent or stable) state is determined by the values of R_1 and E. The diode D serves to limit the negative half-cycle of the input voltage, thereby preventing the break-down of the emitter junctions in the multi-emitter transistor. In many cases, this diode is built into the IC itself. R_2 is part of a parallel diode limiter, R_2D.

By varying the values of R_1 and E as appropriate, we can build two types of multivibrator, differing in the status of their components during the initial (stable) state and, as a consequence, in the polarity of the timing pulse.

Multivibrators Using a Negative-Going Timing Voltage

In this form of multivibrator, the input during the initial state is low, $LG1$ is in the 1 state ($v_{\text{out},1} = V^{\text{H}}$) and $LG2$ is in the 0 state ($v_{\text{out},2} = V^{\text{L}}$) (Fig. 7.9). A switchover to the temporary state occurs at time t_0 due to a negative-going pulse e_2 applied to $LG2$ and bringing about a step rise in $v_{\text{out},2}$, v_A and v_{in}. Then the timing voltage v_A and the input voltage v_{in} gradually decrease as the capacitor recharges. At time t_1, the input voltage reaches a threshold value V'_{th} at which the loop gain exceeds unity and a regeneration process commences to cause $LG1$ and $LG2$ to go back to their initial states.

The initial voltage on the capacitor recovers in two steps. During the first step (interval τ'_{rc}) the limiting diode D is conducting, so the recovery is accompanied by a change in only v_A ($v_{\text{in}} = -e_0$). During the second step, the diode is cut off, so both v_A and v_{in} change at the same time.

Static conditions in the stable state. The manner in which the values of R_1 and E affect the status of $LG1$ is illustrated in Fig. 7.10 which shows the input characteristic (with allowance for the diode D) and the transfer characteristic of the TTL gate with a totem-pole output driver. For $LG1$ to be in the 1 state ($v_{\text{out},1} = V^{\text{H}}$), the operating point must be positioned within portion 1-2 where the inverting transistor ($T1$ in Fig. 5.33) and the diode D are both at cutoff. In the circumstances, the input resistance of the logic gate is determined by the series combination of the emitter junction of the multi-emitter transistor and the resistor R_a in its base lead. Assuming that the voltage across the emitter junction is $e_{\text{EO}} = 0.6$ V, the input characteristic within portion 1-2 may be written

$$i_{\text{in}} = -I^L_{\text{in}} + v_{\text{in}}/r_{\text{in}} \tag{7.20}$$

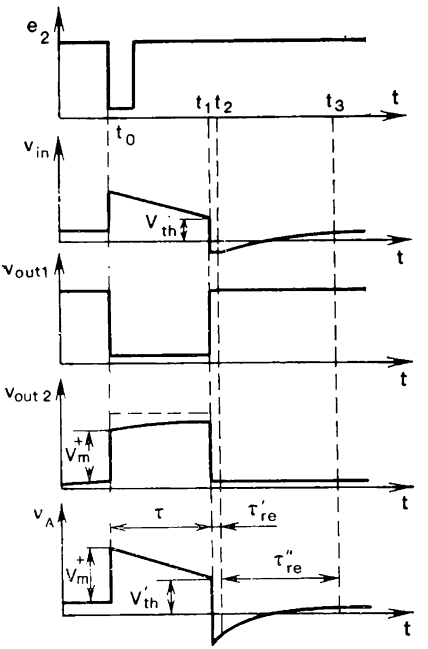


Fig. 7.9

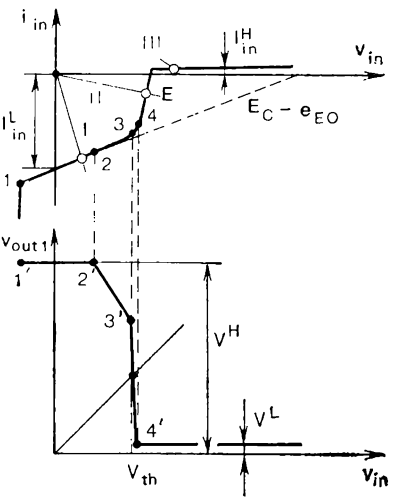


Fig. 7.10

where I_{in}^{L} is the input current as measured at $v_{in} = 0$, $r_{in} = (E_C - e_{EO})/I_{in}^{L}$ is the input resistance, and E_C is the collector supply voltage.

Whatever the status of *LG*1, its steady-state input current (when the capacitive current is zero) can be written

$$i_{in} = (E - v_{in})/R_1 \tag{7.21}$$

In Fig. 7.10, this equation is represented by straight lines I, II and III plotted for various values of E and R_1, the first two lines being plotted for $E = 0$. The intersections between the straight lines and the input characteristic locate the operating point. The steady-state value of input voltage $V_{in,ss,1}$ in the 1 state of *LG*1 (line *I*) can be found by solving Eqs. (7.20) and (7.21) simultaneously, with E set equal to zero:

$$v_{in,\,ss,\,1} = (E_C - e_{E0})/(1 + r_{in}/R_1) \tag{7.22}$$

The position of point *2* where the inverting transistor turns on determines the upper limiting value of R_1:

$$R_{1,\,max} = r_{in}V_{in,\,2}/(E_C - e_{E0} - V_{in,\,2})$$

where $V_{\text{in},2} = e_{\text{EO}} - V^*_{\text{CE(sat)}} \approx 0.5$ V is the value of $V_{\text{in},ss,1}$ at point *2*. For typical values of E_{C} (5 V) and I^{L}_{in} (0.5 to 1.4 mA), $R_{1,\max} = 0.4$ to 1.1 kΩ.

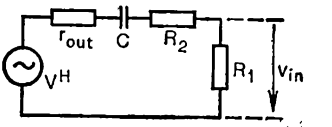


Fig. 7.11

Duration of the temporary state (the interval from t_0 to t_1). The exact duration of this interval cannot be found because at $v_{\text{in}} > V_{\text{th}}$ the input *V-I* characteristic is nonlinear. Because of the low stability of IC logic-gate multivibrators, however, it is legitimate to use for practical purposes an approximate calculation based on the assumption that at $v_{\text{in}} > V_{\text{th}}$ the emitter junction of the multi-emitter transistor is reverse-biased, that is, $i_{\text{in}} = I^{\text{H}}_{\text{in}} \approx 0$. With this assumption, the capacitor-charging circuit can be replaced by the equivalent circuit shown in Fig. 7.11. In view of the exponential character of the transients in the circuit, the duration of the temporary state may be written

$$\tau = \tau_1 \ln \frac{v_{\text{in}}(\infty) - v_{\text{in}}(t_0)}{v_{\text{in}}(\infty) - v_{\text{in}}(t_1)} \tag{7.23}$$

where $\tau_1 = C\,(R_1 + R_2 + r_{\text{out}})$

$v_{\text{in}}(\infty) = 0$

$v_{\text{in}}(t_1) = V_{\text{th}} \approx 1.4$ to 1.5 V

$v_{\text{in}}(t_0) = V_{\text{in, ss, 1}} + (V^{\text{H}} - V^{\text{L}})\, R_1/(R_1 + R_2 + r_{\text{out}})$

r_{out} = output resistance of the logic gate in the 1 state, mainly due to R_3 (see Fig. 5.34) in the collector lead of the output transistor. It can be determined by directly measuring the output *V-I* characteristic of the logic gate in the 1 state.

The condition for the multivibrator to operate normally is stated in terms of the inequality

$$v_{\text{in}}(t_0) > v_{\text{in}}(t_1)$$

By solving it, one can determine the minimum value of R_1. We leave it as an exercise for the reader to derive the applicable relationships.

Recovery (resolving) time (interval from t_1 to t_3). In practice, τ'_{re} may be neglected, because the first step in the recovery process is relatively short. Noting that in the 0 state the output resistance of the second logic gate is practically zero and R_1 is shunted by the input resistance of the first logic gate, τ_{re} may, by analogy with Eqs. (7.10) and (7.11), be written

$$\tau_{re} \approx \tau''_{re} \approx C\,(R_2 + R_1 \parallel r_{\text{in}})\, k_{re}$$

In view of the relatively low accuracy of IC logic-gate multivibrators, we may take a derated value of the exponential decrement

factor ($k_{dec} = 0.8$) for which $k_{re} \approx 1.6$. Yet, the recovery (resolving) time (the interval from t_1 to t_3) for the circuit in question is relatively long and usually exceeds the duration of the temporary state.

The value of R_2 is chosen, assuming that the output transistor of the second logic gate is saturated during the first step of the recovery process. The minimum value of R_2 is given by

$$R_{2,\,min} \approx (V^H - V^L - V_{th} - e_0)/I_{out,\,max}$$

where $I_{out,max}$ is the maximum output current of the second logic gate in the 0 state, equal to NI_{in}^L.

Multivibrators with a Rising Timing Voltage

In the initial state ($V_{in,ss} > V_{th}$), the input of such a multivibrator is high, *LG*1 is in the 0 state ($v_{out,1} = V^L$) and *LG*2 is in the 1 state ($v_{out,2} = V^H$) (Fig. 7.12). A switchover to the temporary state occurs at time t_0 under the action of a negative-going pulse e_1 applied to the input of *LG*1 and causing the logic gates to change state. A step decrease in $v_{out,2}$, v_A, and v_{in} is followed by a re-charge of the capacitor, accompanied by an increase in the timing voltage v_A. During the interval t_0 to t_1, the limiting diode is conducting, so $v_{in} = -e_0$. At time t_1, the diode turns off, and v_{in} and v_A change in value simultaneously during the interval from t_1 to t_2. At time t_2, the input voltage becomes equal to V_{th}, with the result that *LG*1 and *LG*2 jump back to their initial states. The initial voltage on the capacitor is restored during the interval from t_2 to t_3, and the recovery is accompanied by a change in v_{in} and $v_{out,2}$.

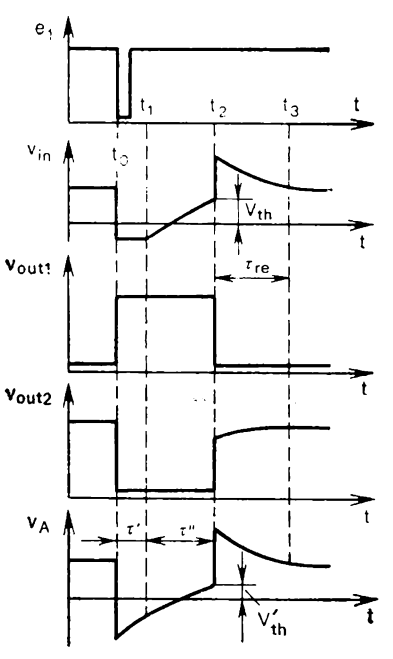


Fig. 7.12

Referring to Fig. 7.10, it is seen that *LG*1 may have a 0 state in any one of two instances. In the first instance occurring when $E = 0$, the straight line described by Eq. (7.21) is labelled II. For this to happen, R_1 must be sufficiently large ($R_1 > R_{1,min} \approx r_{in}V_{th}/(E_C - e_{E0} - V_{th})$. Then, $V_{in,ss,2} \approx V_{th}$. The second instance (straight line III) takes place when $E > V_{th}$. Then $V_{in,ss,3} \approx E$, because I_{in}^H is small.

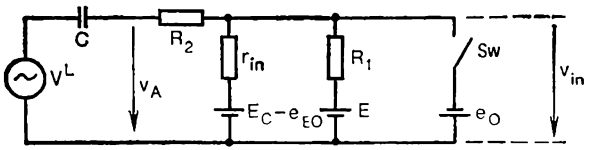


Fig. 7.13

Duration of the temporary state. The manner in which the timing voltage varies can be ascertained from the equivalent circuit of Fig. 7.13. The state of the switch, *Sw*, is determined by the status of the limiting diode. More specifically, it is closed during the first step of the relaxation process, and open during the second. Because the transients during either step are exponential in behaviour, the duration of the temporary state may be written

$$\tau = \tau_1 \ln \frac{v'_A(\infty) - v_A(t_0)}{v'_A(\infty) - v_A(t_1)} + \tau_2 \ln \frac{v''_A(\infty) - v_A(t_1)}{v''_A(\infty) - v_A(t_2)} \tag{7.24a}$$

where

$$\tau_1 = CR_2, \quad \tau_2 = C(R_2 + R_1 \| r_{in})$$

$$v'_A(\infty) = -e_0$$

$$v_A(t_0) = V_{in,\,ss,2(3)} - (V^H - V^L)$$

$$v_A(t_1) = -e_0 - R_2\left(\frac{E + e_0}{R_1} + \frac{E_C + e_0 - e_{EO}}{r_{in}}\right)$$

$$v''_A(\infty) = \frac{Er_{in} + (E_C + e_0)R_1}{r_{in} + R_1}$$

$$v_A(t_2) = V_{th} - R_2\left(\frac{E - V_{th}}{R_1} + \frac{E_C - e_{EO} - V_{th}}{r_{in}}\right)$$

The first step is omitted when

$$R_2 \geqslant (V^H - V^L - V_{in,\,ss,\,2(3)} - e_0) \Big/ \left(\frac{E + e_0}{R_1} + \frac{E_C + e_0 - e_{EO}}{r_{in}}\right)$$

because then $v_A(t_0) \geqslant v_A(t_1)$. As a consequence, the duration of the temporary state is

$$\tau = \tau_2 \ln \frac{v''_A(\infty) - v_A(t_0)}{v''_A(\infty) - v_A(t_2)} \tag{7.24b}$$

Recovery (resolving) time. This time interval is determined by the fact that the capacitor charges up exponentially via the resistors R_2 and R_1 and the output resistance r_{out} of the second logic gate.

At $k_{dec} = 0.8$, $\tau_{re} = 1.6C\ (R_1 + R_2 + r_{out})$. As in a multivibrator with a falling timing voltage, $\tau_{re} > \tau$. To minimize the recovery (resolving) time in a multivibrator with $E > V_{th}$, the resistor R_1 can be shunted by a diode (shown by a dashed line in Fig. 7.8).

Free-Running IC Logic Gate Multivibrators

One way to build a free-running IC logic-gate multivibrator is to make the two coupling circuits of the RS flip-flop capacitive instead of direct. However, such multivibrators are used on a limited scale because they, like all other symmetrical circuits, require a trigger pulse to force them to oscillate, call for a larger count of external circuit components, and present some difficulty for the recovery of the initial voltage on their capacitors during one half-cycle of oscillation.

These difficulties are non-existent in the unsymmetrical free-running multivibrator of Fig. 7.14*a*. In contrast to a triggered multivibrator, the external resistor R_1 is connected to the output of the first logic gate, with the result that the output voltage of the first logic gate acts as the supply voltage E. If $R_1 < R_{1,max}$, both states of the multivibrator are astable, because for $v_{out,1} = V^L \approx 0$, the circuit of Fig. 7.14*a* is equivalent to a multivibrator with a falling timing voltage, whereas for $v_{out,1} = V^H > V_{th}$, it is equivalent to a multivibrator with a rising timing voltage. The timing diagrams of the unsymmetrical multivibrator in question are shown in Fig. 7.14*b*. The values of τ_{01} and τ_{02} can be found, using Eqs. (7.23) and (7.24) on substituting

$$v_{in}(t_0) = V_{th} + (V^H - V^L)\, R_1/(R_1 + R_2 + r_{out})$$

in the former, and

$$v_A(t_0) = V_{th}(1 + R_2/R_1) - (V^H - V^L)$$

in the latter.

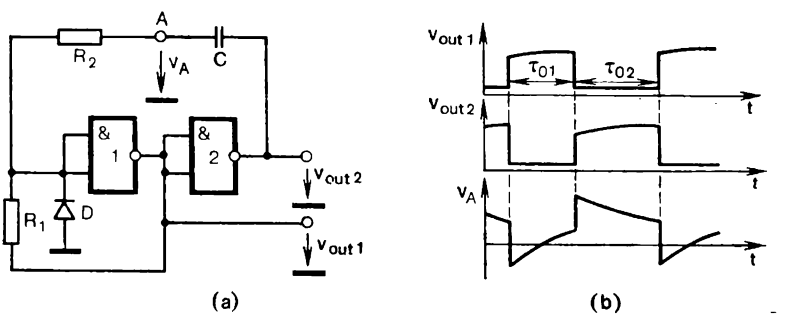


Fig. 7.14

7.5. DIFFERENTIAL OPAMPs AS MULTIVIBRATORS

To act as multivibrators, differential opamps can be arranged in any one of three configurations: with a differentiating *RC*-network in the positive feedback path, with an integrating *RC*-network in the negative feedback path, or with both.

In a multivibrator using a differentiating *RC*-network, the feedback signal exceeds the output signal in amplitude. In a differential stage using bipolar transistors, such a feedback signal would cause the emitter junctions to break down. With the input current properly limited, however, this breakdown could be made reversible, but the relaxation processes would no longer be exponential, the duration of the generated pulses would be reduced, the stability of the multivibrator would be impaired, and its characteristics would be difficult to calculate. These drawbacks limit the use of circuits with differentiating *RC*-networks. The principles underlying their synthesis and calculation are explained in [53].

A multivibrator with an integrating *RC*-network is basically an unsymmetrical multivibrator (see Fig. 6.5) incorporating an additional feedback loop so that it can operate as both a free-running and a monostable multivibrator.

Free-Running Multivibrator with an Integrating RC-Network

Basically, this type of multivibrator (Fig. 7.15*a*) depends for its operation on alternatively charging and discharging its capacitor via the resistor R (see Fig. 7.15*b*).

Suppose that at $t < t_0$, the output voltage is $v = V^{\mathrm{L}}$. This condition will remain so long as $e > E_{\min}$. On varying exponentially, the capacitor voltage tends to $-V^{\mathrm{L}}$. If

$$-V^{\mathrm{L}} < E_{\min} \tag{7.25a}$$

an instant, $t = t_0$, will be reached as the capacitor recharges when e becomes equal to $E_{\min}$, the multivibrator changes state, and, as a consequence, the output voltage becomes equal to $v = V^{\mathrm{H}}$. This condition will remain as long as $e < E_{\max}$. The capacitor voltage tends to V^{H} and with

$$V^{\mathrm{H}} > E_{\max} \tag{7.25b}$$

at time $t = t_1$ it reaches the level of $E_{\max}$ at which the circuit resets. The events described above repeat themselves periodically. The period of oscillation is the sum of two half-cycles ($T = \tau_{01} + \tau_{02}$)

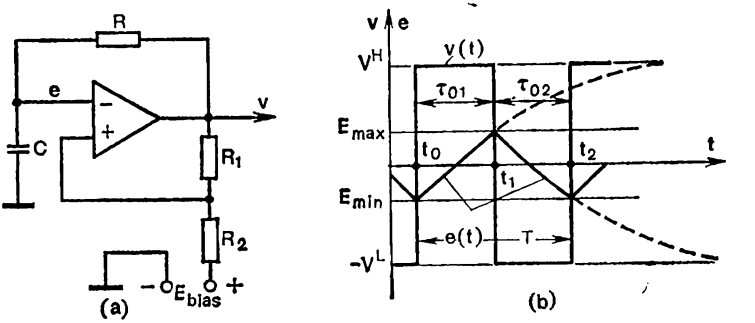


Fig. 7.15

during which the timing voltage e completes a full cycle of alternations between E_{max} and E_{min}. Accordingly, we may write

$$T = \tau \left(\ln \frac{V^H - E_{min}}{V^H - E_{max}} + \ln \frac{V^L + E_{max}}{V^L + E_{min}} \right)$$

where $\tau = RC$. Using Eq. (6.8) where $\gamma = R_2 (R_1 + R_2)$, we get

$$T = \tau \ln \left[1 + \frac{\Delta E}{(1-\gamma)(V^H - E_{bias})} \right] + \tau \ln \left[1 + \frac{\Delta E}{(1-\gamma)(V^L + E_{bias})} \right]$$

In practice, the second terms of the logarithms are usually small, therefore

$$T \approx \tau \Delta E \Delta V / [(1-\gamma)(V^H - E_{bias})(V^L + E_{bias})]$$

where $\Delta V = V^H + V^L$ is the amplitude of the signal at the amplifier output.

The pulse period to pulse duration ratio (the reciprocal pulse duty) for the multivibrator in question is given by

$$Q_p = T/\tau_{01} \approx 1 + (V^H - E_{bias})/(V^L + E_{bias})$$

If the circuit is symmetrical, $V^H = V^L$, $E_{bias} = 0$, and $\Delta V = 2V^H$, then

$$T = RC\,[4\gamma/(1-\gamma)] \text{ and } Q_p =$$

In practice, the symmetry is upset because $V^H \neq V^L$. The situation can be improved through the use of silicon Zener diodes (Fig. 7.16). Additionally, they will make the pulse repetition period more stable.

Large values of Q_p can be obtained by shunting R with a diode-resistor network (D_3R' in Fig. 7.17) so as to upset the symmetry of the circuit. Then the capacitor will charge via R and R' connected in parallel, whereas during discharge the diode will be nonconducting.

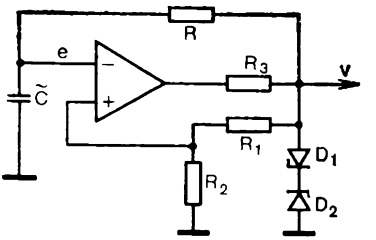


Fig. 7.16

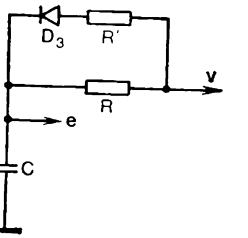


Fig. 7.17

The desired pulse period to pulse duration ratio can be obtained by suitably varying the values of R and R'. It is to be noted that the minimum values that may be chosen for R, R' and $R_1 + R_2$ are limited by the load capacity of the amplifier.

One-Shot Multivibrator with an Integrating RC-Network

The above multivibrator can be re-arranged into a one-shot by setting E_{bias} such that one of the inequalities (7.25) is violated. As a result, one of the temporary states will become a stable one. This approach is applicable in cases where the range of input voltages of the differential opamp is wider than the range of output voltage. This is not the case with most of the integrated circuits in use. Therefore, it is more customary to obtain one-shot operation by interposing a diode as shown in Fig. 7.18*a* so as to limit the time during which the capacitor re-charges. Such a multivibrator will be in its stable state when the output of the differentiating opamp goes high. Under steady-state conditions, the diode is conducting and the voltage e is clamped to e_0 equal to the voltage drop across the conducting p-n junction. For this state to be stable, it is essential

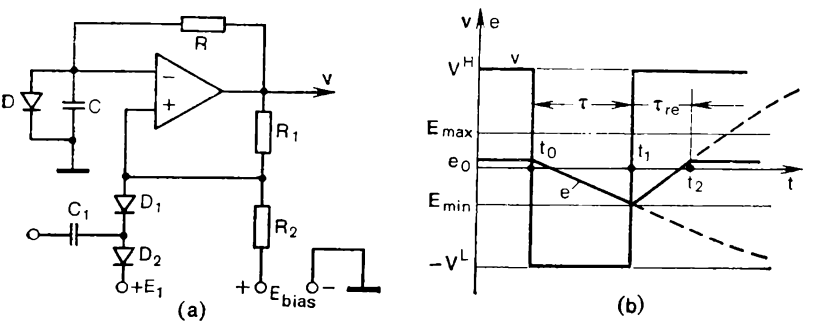


Fig. 7.18

that $e_0 < E_{max}$, which is ensured, for example, when $E_{bias} = 0$ and $\gamma > e_0 V^H$.

A negative-going pulse applied to the noninverting input of the differentiating opamp causes the multivibrator to go from its stable to its temporary state with a low output voltage (Fig. 7.18*b*). When this happens, the capacitor begins to discharge until the time t_1 when $e = E_{min}$. The circuit goes back to its original state, and the original voltage is restored on the capacitor. The recovery is complete by the instant when the diode turns on.

Duration of the generated pulse. During the interval from t_0 to t_1 the timing voltage varies exponentially

$$e(t) = (e_0 + V^L) \exp[-(t - t_0)/RC] - V^L$$

From the equation

$$e(t_1) = E_{min}$$

the duration of the generated pulse is found to be

$$\tau = t_1 - t_0 = RC \ln (e_0 + V^L)/(E_{min} + V^L)$$

$$= RC \ln \left(1 + \frac{e_0 - E_{min}}{E_{min} + V^L}\right)$$

If $E = 0$, $e_0 \ll V^L$, and $\gamma \ll 1$, then

$$\tau = RC \ln \left(1 + \frac{\gamma}{1+\gamma}\right) \approx RC \ln (1 + \gamma) \approx \gamma RC$$

Recovery (resolving) time. During the recovery interval, $\tau_{re} = t_2 - t_1$, exponential variations in $e(t)$ can be written as

$$e(t) = (E_{min} - V^H) \exp(-t/RC) + V^H$$

On substituting $e(t_2) = e_0$, we obtain

$$\tau_{re} = RC \ln \frac{V^H - E_{min}}{V^H - e_0} = RC \ln \left(1 + \frac{e_0 - E_{min}}{V^H - e_0}\right) \qquad (7.26)$$

The recovery (resolving) time can be reduced by shunting R with a diode network (Fig. 7.16). Then, R in Eq. (7.26) must be replaced by $RR'/(R + R')$.

REVIEW QUESTIONS AND PROBLEMS

7.1. Name the states which a one-shot and a free-running multivibrator can have.

7.2. Describe the relaxation process.

7.3. Plot a timing diagram for the timing voltage of a symmetrical multivibrator using *p-n-p* transistors and integrated *n-p-n* transistors, an unsymmetrical free-running multivibrator made with IC logic gates, and an opamp free-running multivibrator.

7.4. In the directly-coupled one-shot multivibrator of Fig. 7.1*a*.

transistor $T1$ has been shunted by a resistor whose value is R_{C1} ($\gamma = 1$). Find how this has reduced the duration of the output pulse.

7.5. In the directly coupled one-shot multivibrator of Fig. 7.1*a*, the wiper of R_{C1} is set in the middle position ($\gamma = 0.5$). Plot the waveform of the pulse in the collector of $T1$. Derive the relation connecting τ_{re} and γ.

7.6. In a directly coupled one-shot multivibrator operating at $E_C = 12$ V, the amplitude of pulses in the base of $T2$ is 3 V. Using the ideal transistor model, find k_τ.

7.7. Name the elements which should be added to a symmetrical multivibrator so as to increase the output pulse period to pulse duration ratio (the reciprocal output pulse duty).

7.8. What is the function of diodes $D1$ and $D2$ in the К2ГФ182 IC multivibrator shown in Fig. 7.5*a*?

7.9. How does the diode $D4$ in the К2ГФ182 IC multivibrator of Fig. 7.5*a* affect the waveform of the trigger voltage of the multivibrator?

7.10. How should the external components be connected in the К2ГФ182 IC multivibrator of Fig. 7.5*a* so as to make the multivibrator free-running?

7.11. Explain the function of R_{01} and R_{02} in the К2ГФ192 IC multivibrator of Fig. 7.7.

7.12. Draw the circuit and calculate the circuit parameters of an IC logic-gate multivibrator with a falling trigger voltage. The output pulses are 10 μs in duration, and the multivibrator is made with TTL gates with a totem-pole output driver, for which $E_C = 5$ V, $V^H = 3.8$ V, $V^L = 0.1$ V, $N = 8$, $I_{in}^L = 1$ mA, $r_{out} = 150\ \Omega$. It is assumed that each output of the multivibrator can be loaded into two identical logic gates. Calculate the maximum repetition rate of trigger pulses.

7.13. Work out Problem 7.12, using the circuit with a rising trigger voltage.

7.14. Using the TTL gates defined in Problem 7.12, design a free-running multivibrator generating output pulses with a repetition rate of 300 kHz. It is assumed that each output of the multivibrator can be loaded into no more than two identical logic gates.

7.15. Calculate the values of the external components for the one-shot multivibrator of Fig. 7.18*a*, generating pulses with a duration of 1 ms and a pulse repetition rate of not over 0.5 kHz, if the amplifier is a 1УТ401Б integrated circuit for which $V^H = 10$ V and $V^L = -7$ V across $R_L = 5$ kΩ. The input currents of the amplifier may be neglected in calculations.

7.16. How should the circuit of the multivibrator in Problem 7.15 be modified if the pulse repetition rate should be raised to 0.9 kHz? If the output pulses should be of positive polarity?

8 BLOCKING OSCILLATORS

8.1. GENERAL

A blocking oscillator is a relaxation oscillator generating square pulses, consisting essentially of a single-stage amplifier with a pulse transformer in the positive feedback path.

As regards the stability of the output waveform and period of oscillation, blocking oscillators are usually inferior to multivibrators, but they offer advantages of their own. For one thing, the amplifier transistor in a blocking oscillator draws current during only a small fraction of the entire cycle, so that the output pulses may have a high peak power, a high pulse period to pulse duration ratio ($Q_L \gg 10$), and a low average power. For another, the pulse transformer can supply pulses of any polarity and with an amplitude which may exceed the supply voltage, the output resistance is low, and the circuit is highly reliable because it uses a small count of circuit components. The principal drawback of blocking oscillators is the need for a pulse transformer which cannot be microminiaturized.

At present, blocking oscillators are used as drivers and modulators for r.f. pulse generators, clock pulse generators for magnetic memories, one-bit stages in storage counters, and low-to-high d.c. voltage converters. Like any other relaxation oscillators (such as multivibrators and phantastrons), they can be triggered, free-running, or synchronized, and operate as frequency dividers.

8.2. COLLECTOR-COUPLED BLOCKING OSCILLATOR

The basic circuit of a blocking oscillator in which the collector is coupled to the base by a pulse transformer is shown in Fig. 8.1*a*. Because the base resistor R is returned to the $-E_C$ supply, the blocking oscillator is free-running. For triggered operation, R must be returned to the $+E_B$ supply.

Let us trace the operation of the blocking oscillator, starting at time $t < t_0$ when the timing capacitor C charged during the previous pulse ceases to discharge. The waveforms of the voltages on the base and collector of the transistor, the capacitor C, and the load resistor R_L connected in the tertiary winding of the transformer are shown in Fig. 8.1*b*.

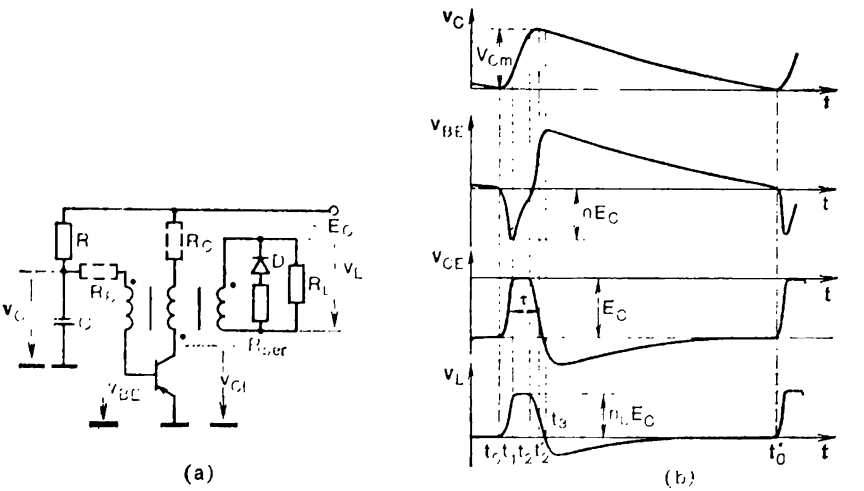


Fig. 8.1

The timing capacitor discharges at a relatively slow rate (this is a relaxation process). Therefore, practically no emf is induced in the transformer windings, the base voltage is the same as the capacitor voltage, and the transistor is at cutoff. The pulse leading edge begins to be formed at $t = t_0$ when the base is pulled up to the cut-in voltage. The collector current completes the positive feedback path through the collector and base windings connected in opposition (see Fig. 8.1*a*). Soon after the controlled currents begin to flow through the transistor, the base current gain becomes sufficiently large for the feedback loop gain to be $K_{FB} > 1$, so that the transistor currents build up cumulatively and at a high rate—the leading edge now rises regeneratively.

The circuit parameters are chosen such that the leading edge of the output pulse is fully formed at the instant when the transistor moves into the saturated region. Because the input resistance determining the equivalent collector load is small, the switchover to saturation occurs at heavy currents hundreds or even thousands of times the safe static limits. If the average power dissipation is likely to exceed the safe limit, these currents are often limited by inserting resistors R_C and R_B shown by dashed lines in Fig. 8.1*a*. Also, R_C can sometimes be used as a source of voltage pulses free from backswing (tail), and R_B serves to stabilize the pulse duration in spite of variations in r_{in} due to changes in the base bulk resistance (base resistance modulation).

At the end of the rise time (at time $t = t_1$), the base current in a blocking oscillator using a transformer with an optimal transform-

ation ratio will always exceed the collector current. Because of this, the transistor is driven hard into saturation. Removal of the minority-carrier excess charge, occupying an interval (the storage time) longer than the pulse rise time, defines the duration (t_1 to t_2) of the pulse top, which is close to the pulse duration τ.

Over the pulse duration, the capacitor is charged by the falling base current determining the change in the total minority-carrier charge. In contrast to resistor-coupled switches used in flip-flops and multivibrators, the collector current of a blocking oscillator at saturation does not remain constant. It is the sum of the direct component, the rising magnetizing current of the transformer, and the decreasing base current. Therefore, the collector current and, as a consequence, the related minority-carrier charge may rise or fall, depending on the rate of change of the transformer and base currents.

At the end of the pulse duration (at time $t = t_2$), the transistor moves out of saturation, and the minority-carrier charge rapidly decreases until it becomes equal to its limiting value. With the transistor moving into the active region, the positive feedback path is completed in the same manner as during the leading edge of the pulse, and the major portion of the trailing edge begins to be formed, with the transistor currents falling at a progressively increasing rate. At time t_2', the emitter junction of the transistor is reverse-biased, and the transistor is driven to cutoff. This produces the remaining portion of the trailing edge (t_2' to t_3) and the backswing (t_3 to t_0') in the collector voltage pulse, owing to the leakage of magnetic energy stored in the transformer core while the pulse top was formed.

The interval between pulses, T-τ, lasts so long as the capacitor voltage v_C about equal to v_{BE} decreases as the timing capacitor C discharges or, rather charges via R from $-E_C$ to the value from which we began our discussion.

Conditions for operation of a blocking oscillator. The key requirement for a blocking oscillator to operate normally is that the switchover must be regenerative in nature. For this to happen, the rise and fall times of the output pulse must be such that

$$K_{FB} \approx nK_{eq} > 1 \tag{8.1a}$$

where $n = w_B/w_C$, $K_{eq} = \beta\,(R_L' \parallel r_{in}')/r_{in}$, $R_L' = R_L/n_L^2$, $r_{in}' = r_{in}/n^2$, $n_L = w_L/w_C$, w_C, w_B and w_L are the number of turns in the collector, base and load windings of the transformer, respectively.

In deriving Eq. (8.1a), the capacitor C was visualized as a short-circuit for fast variations in voltage during the rise and fall times. From Eq. (8.1a) it follows that the blocking oscillator will operate normally if the load is

$$R_L' > R_{L,\,min}' \tag{8.1b}$$

where $R_{L,min}' = r_{in}/[n\,(\beta - 1)]$.

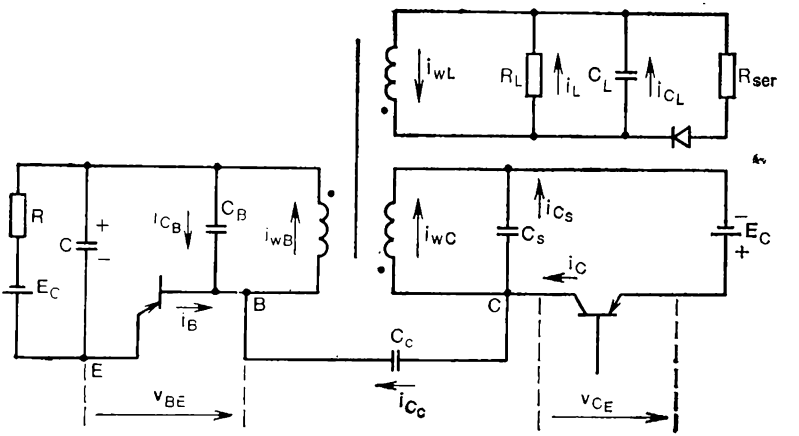


Fig. 8.2

If the load resistance is high so that $R'_L \gg r'_{in}$, it is required that

$$\beta/n > 1 \tag{8.1c}$$

If the base circuit contains R_B, r'_{in} in Eq. (8.1a) must be replaced with $R_{in} = (R_B + r_{in})/n^2$. It follows from Eqs. (8.1) that blocking oscillators have, as already noted, a high loading capacity.

Basic equations for blocking oscillators. Let us rearrange the circuit of Fig. 8.1*a* as shown in Fig. 8.2. From inspection of this circuit, we may write

$$\begin{aligned} E_C &= v_{CE} + (v_{BE} + v_C)/n \\ i_{wL} &= i_L + i_{C_L} \\ i_{wB} &= i_B + i_{C_C} + i_{C_B} \\ i_{wC} &= i_C - i_{C_s} - i_{C_C} \end{aligned} \tag{8.2a}$$

To transfer the parameters of Fig. 8.2 into the collector circuit, it is necessary to use the equation of balance for the ampere-turns of the transformer:

$$w_C i_\mu = w_C i_{wC} - w_B i_{wB} - w_L i_{wL} \tag{8.2b}$$

In view of Eq. (8.2b), we may write Eqs. (8.2a) as

$$i_C = i_\mu + i'_B + i'_L + i'_C \tag{8.2c}$$

where

$$\begin{aligned} i'_B &= n i_B \\ i'_L &= n_L i_L \\ i'_C &= i_{C_s} + i_{C_C}(1+n) + n i_{C_B} + n_L i_{C_L} \end{aligned}$$

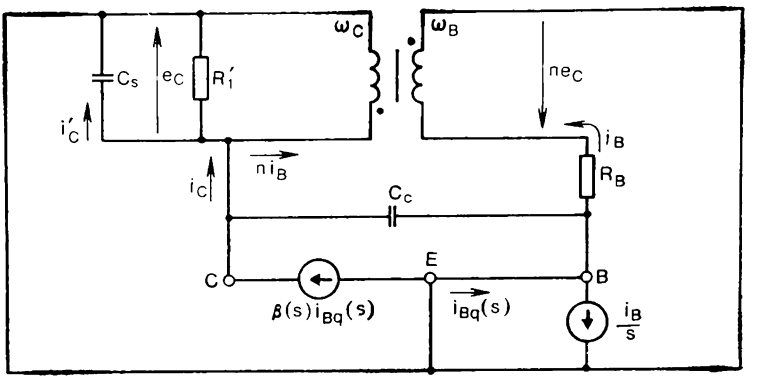


Fig. 8.3

The first line of Eqs. (8.2a) is the voltage equation, and Eq. (8.2c) is the current equation. They are the basic equations of a blocking oscillator, used when finding the circuit parameters referred to the collector.

Using Eqs. (8.2a) and (8.2c), we can draw equivalent circuits which permit analysis of the fast processes—the shaping of the leading and trailing edges (the rise and fall times), and also the relatively slow processes—the shaping of the pulse top (pulse width) and the recovery of the initial state.

Leading edge. For the duration of the rise time (the leading edge), the transistor is in the active region. To a first approximation, we may deem that for the duration of the rise time, the magnetizing current of the transformer and the voltage on the timing capacitor have no time to depart from their initial (zero) values. Also, it should be remembered that the capacitive current has a marked effect on the shape of the leading edge. In view of these remarks, the a.c. component existing over the pulse rise time can be analyzed, using the equivalent circuit of Fig. 8.3. Here, the transistor is represented by the simplified model discussed in Chap. 3. The control current generator I_B/s simulates the effect of fluctuations (in the free-running mode) or trigger pulses (in the triggered mode). The leakage inductance L_{lk} is assumed to be negligibly small, as it usually the case in pulse transformers.

The stray capacitance C_s is connected to the capacitive current i_C' in the usual manner:

$$C_s = i_C'/(de_C/dt)$$

Assuming that the transformer in the circuit of Fig. 8.3 is ideal, we may write the Laplace transform of the change of collector volt-

age as

$$v_{CE\sim}(s) = \frac{I_B}{s} \frac{\beta R'_L}{\tau_\beta \tau_s^*} \frac{1}{(s-s_1)(s-s_2)} \tag{8.3a}$$

where

$$\tau_s^* = R'_L (C_s + C_c)$$

and the roots of the denominator are found from the characteristic equation

$$s^2 + s \frac{R_B(\tau_\beta + \tau_s^*) + n^2 R'_L \tau_\beta}{R_B \tau_\beta \tau_s^*} + \frac{R_B + nR'_L(n-\beta)}{R_B \tau_\beta \tau_s^*} = 0 \tag{8.3b}$$

Hence,

$$s_{1,2} = -\frac{R_B(\tau_\beta + \tau_s^*) + n^2 R'_L \tau_\beta}{2R_B \tau_\beta \tau_s^*} \times \left[1 \mp \sqrt{1 - \frac{4R_B \tau_\beta \tau_s^* [R_B + nR'_L(n-\beta)]}{[R_B(\tau_\beta + \tau_s^*) + n^2 R'_L \tau_\beta]^2}}\right]$$

In practice, the second term of the radicand is usually markedly smaller than unity. Therefore,

$$s_1 \approx -\frac{1 + R'_L(1-\beta/n)/R'_B}{\tau_s^* + \tau_\beta(1 + R'_L/R'_B)} \tag{8.3c}$$

$$s_2 \approx -[1/C_s^*(R'_L \parallel R'_B) + 1/\tau_\beta] \tag{8.3d}$$

where $C_s^* = C_s + C_c$.

It follows from Eq. (8.3c) that the cumulative build-up of voltage (the triggering of the blocking oscillator) can take place only if the root s_1 is positive, which corresponds to the condition

$$\beta/n > 1 + R'_B/R_L \tag{8.4}$$

Equation (8.4) is identical with Eq. (8.1) defining the condition for the correct operation of a blocking oscillator, with the exception that in the adopted transistor model the input resistance, r_{in}, is R_B. When the inequality (8.4) is satisfied well, the Laplace transform (8.3a) can be simplified:

$$v_{CE\sim}(s) \approx -\frac{I_B R_B}{sn} \frac{1}{(1+s\tau_2)(1-s\tau_1)} \tag{8.5a}$$

where

$$\tau_1 \approx \frac{n}{\beta}\left[C_s^* R'_B + \frac{\tau_\beta(1+R'_B)}{R'_L}\right] \tag{8.5b}$$

$$\tau_2 = \tau_\beta \parallel [C_s^*(R'_L \parallel R'_B)] \tag{8.5c}$$

The original time function for the Laplace transform (8.5a) is

$$v_{CE\sim}(t) = \frac{I_B R_B}{n}\left[\frac{\exp(t/\tau_1)}{1+\tau_2/\tau_1} + \frac{\exp(-t/\tau_2)}{1+\tau_1/\tau_2} - 1\right] \tag{8.5d}$$

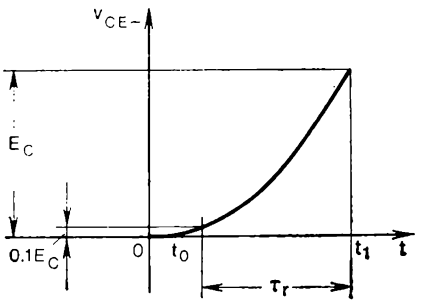


Fig. 8.4

A plot of this function for a small initial disturbance, $I_B R_B/n \ll 0.1E_C$, is given in Fig. 8.4. In this case, the pulse rise time, τ_r, is taken to be the interval from t_0 to t_1, where $v_{CE\sim}(t_0) = 0.1E_C$ and $v_{CE\sim}(t_1) = E_C$. The interval from zero to t_0 may be treated as the pulse delay. Because we have assumed that the initial disturbance is small, $\exp(t_0/\tau_1) \gg 1$, then

$$v_{CE\sim}(t) \approx \frac{I_B R_B}{n} \frac{\exp(t/\tau_1)}{1+\tau_2/\tau_1} \tag{8.6}$$

Thus, under the assumptions made, the initial disturbance will only affect the pulse delay.

On substituting $v_{CE\sim}(t_0) = 0.1E_C$ and $v_{CE\sim}(t_1) = E_C$ in Eq. (8.6), we find the pulse rise time (the duration of the leading edge) to be

$$\tau_r = t_1 - t_0 \approx 2.3\tau_1 \tag{8.7}$$

It follows from Eqs. (8.7) and (8.5b) that there is an optimal value of the transformation ratio for which the pulse rise time is a minimum:

$$n_{opt} = \sqrt{R_B\left(\frac{C_s+C_c}{\tau_\beta} + 1/R'_L\right)} \tag{8.8}$$

For typical parameter values, $n_{opt} \leqslant 0.5$.

It should be remembered that the transformation ratio falls to its minimum very gradually. Therefore, assuming that the pulse transformer is similar to that examined in Chap. 2, the value of n is often chosen to be close to unity, including the value of R_C.

Pulse top. For the interval (t_1 to t_2) occupied by the pulse top (see Fig. 8.1*b*), the equivalent circuit of Fig. 8.2 can be re-arranged as shown in Fig. 8.5. Here, the output and input circuits of the transistor are respectively represented by the output and input resistances at saturation, and the capacitive current flowing through the stray capacitance C_s is assumed to be negligible because the voltages across the transformer windings vary at a very slow rate. The equivalent circuit is also extended to include R_B which, as already noted, minimizes the effect of the input-circuit parameters on the pulse duration and provides a means for controlling the pulse duration. Furthermore, the equivalent circuit does not contain the relative high-valued timing resistor R.

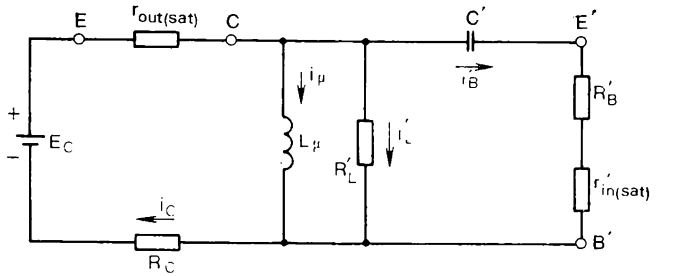


Fig. 8.5

To simplify the analysis, R_C and $r_{out(sat)}$ will be deemed shorted out, because practical blocking oscillators usually satisfy the inequality $L_\mu/(R_C + r_{out(sat)}) \gg \tau$ well. The total minority-carrier charge Q and the limiting minority-carrier charge Q_{lim} can be found from the following relations:

$$\tau_{sat}\,(dQ/dt) + Q = \tau_{sat} i_B\,(t) \tag{8.9a}$$

$$Q_{lim}\,(t) = \tau_B i_C\,(t) = \tau_B\,[i_\mu\,(t) + i'_L\,(t) + i'_B\,(t)] \tag{8.9b}$$

where

$$\tau_B = \tau_\beta/\beta$$

Under the assumption that R_C and $r_{out(sat)}$ are shorted out, the currents are found to be

$$\begin{aligned} i_\mu &= E_C t/L_\mu \\ i'_L &= E_C/R'_L \\ i'_B\,(t) &= I'_{B,\,m} \exp\,(-t/\tau_0) \\ i_B\,(t) &= i'_B\,(t)/n \end{aligned} \tag{8.10}$$

where $I'_{B,m} = E_C/(R'_B + r'_{in(sat)}$ and $\tau_0 = C\,(R_B + r_{in(sat)})$.

As usual, time in Eqs. (8.9) and (8.10) is reckoned from the start of the interval t_1 to t_2. On substituting Eq. (8.10) in Eq. (8.9), we obtain

$$Q\,(t) = I_{B,\,m}\frac{\tau_0\tau_{sat}}{\tau_0 - \tau_{sat}}$$

$$\times\,[\exp\,(-t/\tau_0) - \exp\,(-t/\tau_{sat})] + Q\,(0)\exp\,(-t/\tau_{sat}) \tag{8.11a}$$

$$Q_{lim}\,(t) = \tau_B\,[E_C t/L_\mu + E_C/R'_L + I'_{B,\,m}\exp\,(-t/\tau_0)] \tag{8.11b}$$

where

$$Q\,(0) \equiv Q\,(t_1) = \tau_B i_C\,(t_1) \approx \tau_B E_C/[R'_L \parallel (R'_{in(sat)} + R'_B)]$$

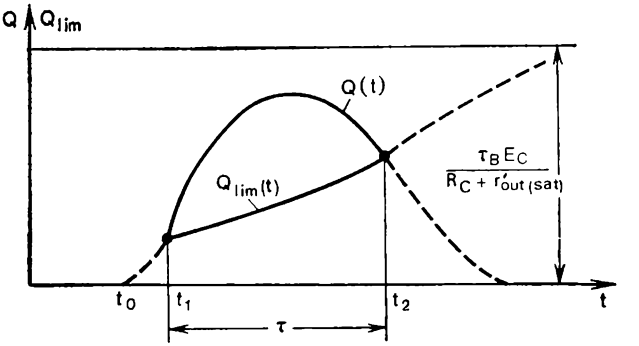


Fig. 8.6

Equations (8.11) are illustrated in Fig. 8.6. The termination of the pulse duration ($t = t_2$) can be found from Eq. (8.11) on substituting

$$Q(\tau) = Q_{lim}(\tau) \tag{8.12}$$

The transcendental equation (8.12) can be solved numerically on a computer. Analytically, it is advisable to examine the limiting cases typical of blocking oscillators with a short and a long pulse duration.

For blocking oscillators with a short pulse duration ($\tau = 1$ to 0.01 μs), the time-setting capacitance is usually small so that

$$4\tau_0 < \tau \ll \tau_{sat} \tag{8.13}$$

In this case, the exponential terms in Eq. (8.11) related to the fall of the base current at the end of the pulse width may be deemed negligibly small. Also, the exponential terms with the time constant τ_{sat}, defining variations in the total charge, may be approximated by the linear terms of a power series. Thus,

$$Q(\tau) \approx [\tau_0 I_{B,m} + Q(0)](1 - \tau/\tau_{sat}) \tag{8.14a}$$

$$Q_{lim}(\tau) \approx \tau_B [E_C \tau/L_\mu + E_C/R'_L] \tag{8.14b}$$

In Fig. 8.7*a*, Eqs. (8.14) are illustrated by solid lines. The dashed lines represent the initial portions of the functions $Q(t)$ and $Q_{lim}(t)$ during the decay of the recombination charge and the limiting charge associated with the referred base current.

From Eqs. (8.12) and (8.14), the pulse duration is found to be

$$\begin{aligned}\tau = {} & [L_\mu (1 + \tau_0/n\tau_B)/(R'_B + r'_{in(sat)})] \\ & \times \{1 + L_\mu [1 + \tau_0/(n\tau_B) + (R'_B + r'_{in(sat)})/R'_L] \\ & \times [\tau_{sat} (R'_B + r'_{in(sat)})]^{-1}\}^{-1}\end{aligned} \tag{8.15}$$

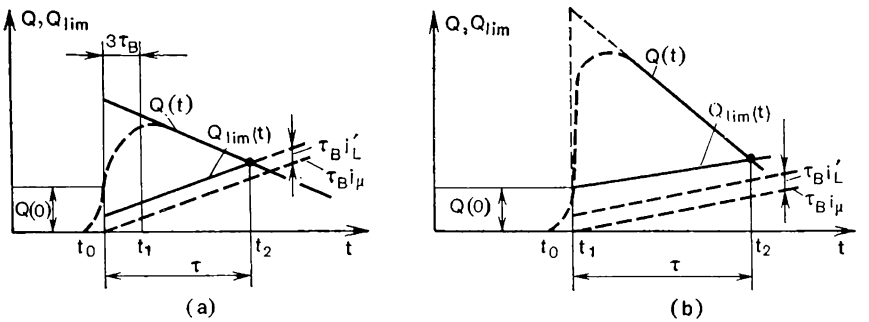


Fig. 8.7

It follows from Eq. (8.15) that in the case in question the pulse duration markedly depends on the time constants τ_B and τ_{sat} characterizing the inertia of the transistor. It may be added that this dependence is more pronounced for high-frequency transistors.

For blocking oscillators with a long pulse duration ($\tau \geqslant 10\mu s$) the time-setting capacitance is high, so that

$$4\tau_{sat} < r \ll \tau_0 \tag{8.16}$$

Now the exponential terms in Eqs. (8.11) associated with the excess minority-carrier charge storage may be neglected at time $t = \tau$, and the exponential terms with the time constant τ_0 can be approximated by the linear terms of a power series:

$$Q(\tau) \approx I_{B,m}\tau_{sat}(1-\tau/\tau_0) \tag{8.17a}$$

$$Q_{lim}(\tau) \approx \tau_B[E_C\tau/L_\mu + E_C/R'_L + I'_{B,m}(1-\tau/\tau_0)] \tag{8.17b}$$

In Fig. 8.7b, Eqs. (8.17) are illustrated by solid lines. Using Eqs. (8.12) and (8.17a), we get

$$\tau = \frac{L_\mu}{R'_B + r'_{in(sat)}} \times \frac{a\beta/n - 1 - (R'_B + r'_{in(sat)})/R'_L}{1 + L_\mu(a\beta/n-1)/[\tau_0(R'_B + r'_{in(sat)})]} \tag{8.18}$$

where $a = \tau_{sat}/\tau_\beta$.

It follows from Eq. (8.18) that in blocking oscillators generating pulses of long duration, the pulse duration is mainly affected by the common-base current transfer ratio, β.

The limiting case of the blocking oscillator in question is the so-called *L*-oscillator—a blocking oscillator without a time-setting

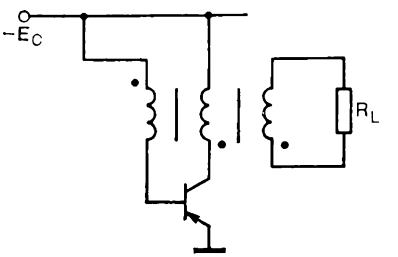


Fig. 8.8

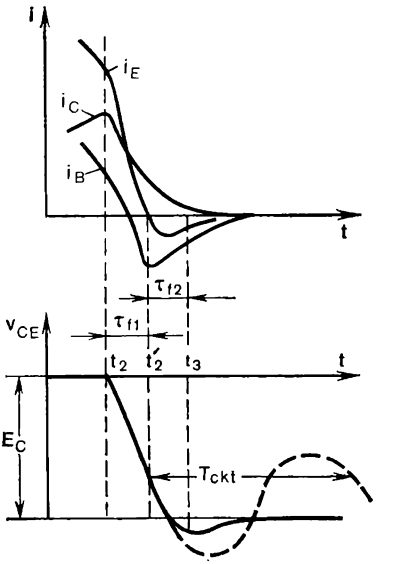


Fig. 8.9

RC-network (Fig. 8.8), which is the key element in low-to-high voltage converters [14]. On setting $\tau_0 = \infty$ in Eq. (8.18),

$$\tau = \frac{nL_\mu a\beta}{R_B + r_{in(sat)}} \left[1 - (n/a\beta) \left(1 + \frac{R'_B + r'_{in(sat)}}{R'_L} \right) \right] \qquad (8.19)$$

Trailing edge. The time interval from t_2 to t_3 occupied by the trailing edge, or the pulse fall time, τ_f, may be treated as the sum of two subintervals: from t_2 to $t'_2 = \tau_{f1}$ and from t'_2 to $t_3 = \tau_{f2}$ (Fig. 8.9). During the first subinterval, which is the fall proper, the transistor is in the active region, so the currents and voltages change cumulatively at a progressively increasing rate, as they do during the rise time. The fall portion of the trailing edge is proportional in duration to the time constant τ_1 and can be minimized by optimizing the transformation ratio, n_{opt}. The remaining or decay portion of the trailing edge is largely shaped with the transistor at cutoff (leaving out in a first approximation the dynamic current cutoff examined in Chap. 3). During the fall portion of the trailing edge, the magnetizing current of the transformer remains nearly constant and at the end of the pulse top it is equal to $i_\mu (t'_2) \approx i_\mu (t_2) = E_C\tau/L_\mu$. During the decay portion, the magnetizing current begins to diminish.

The equivalent circuit that reflects this decrease in the magnetizing current is shown in Fig. 8.10. During the interval from t'_2 to t_3

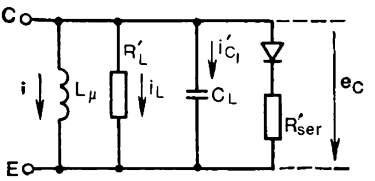


Fig. 8.10

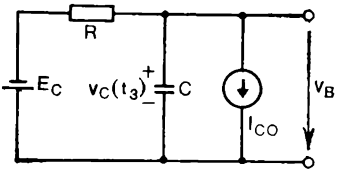


Fig. 8.11

the diode is OFF, and the equivalent resonant circuit $L_\mu/R'_L/C_L$ has an oscillatory response, because $R'_L > R_{cr} = 0.5\sqrt{L_\mu/C_L}$. Approximately, it may be taken that the duration of the minor portion of the trailing edge, or the decay time τ_{f2}, is

$$\tau_{f2} \approx 0.25T_C \approx 0.5\pi\sqrt{L_\mu C_L} \tag{8.20}$$

Which of the two portions, the fall or the decay, will prevail depends on the relative values of the circuit components.

Recovery time of a blocking oscillator. The recovery of the blocking oscillator shown in Fig. 8.1 is essentially controlled by two processes: the dissipation of magnetic energy stored in the transformer (the decrease in the magnetizing current) and the discharge (or, more accurately, recharge) of the time-setting capacitor. The equivalent circuit applying to the former is shown in Fig. 8.10 where the diode turns ON at $t \geqslant t_3$, and the resultant shunting resistance of the resonant circuit $R_{sh} = R'_L + R'_{ser}$, where $R'_{ser} = R_{ser}/n_L^2$, decreases. To cut down the damping time, R'_{ser} is usually chosen such that the resonant circuit has a critical or slightly oscillatory transient response, in accordance with the relation $R_{sh} \approx R_{cr} = 0.5\sqrt{L_\mu/C_L}$. In the circumstances, the duration and amplitude of the backswing, τ_{bs} and $V_{m,bs}$, are given by

$$\tau_{bs} \approx 2.4\sqrt{L_\mu C_s} \text{ and } V_{m,bs} \approx i_\mu(t_3) \times 0.37\sqrt{L_\mu/C_s} \tag{8.21}$$

The second process, the discharge of the time-setting capacitor charged during the pulse top, via the time-setting resistor (the relaxation process) is illustrated in the equivalent circuit of Fig. 8.11. Assuming that a cycle of oscillation begins (the transistor turns on) at $v_B(\tau_{re}) \approx 0$, we obtain from Fig. 8.11:

$$\tau_{relax} \approx \tau_{re} \ln[1 + v_C(t_3)/(E_C + I_{CO}R)] \tag{8.22}$$

where $\tau_{re} = RC$.

The initial voltage across the capacitor is

$$v_C(t_3) \approx v_C(t_2) - \Delta V_C$$

where $v_C(t_2) = nv_C(t_2) = nE_C[1 - \exp(-\tau/\tau_0)]$, and $\Delta V_C \approx Q(t_2)/C$ is the capacitor discharge during the pulse fall time.

To minimize the temperature instability due to I_{CO}, the value of R should be chosen such that $I_{CO,\,max}\,R \ll E_C$.

In blocking oscillators with a time-setting RC-network, the capacitor discharge is the prevailing process, so that the recovery time is equal to the relaxation time, $\tau_{re} = \tau_{relax}$. Thus, for the circuit of Fig. 8.1, the period of oscillation is

$$T = \tau + \tau_{relax} \approx \tau_{relax} \tag{8.23}$$

where it is recognized that $\tau_{relax} \gg \tau$.

8.3. MODIFIED BLOCKING OSCILLATORS

The basic blocking-oscillator circuit examined above can be modified so as to improve the stability of pulse duration and oscillation period despite variations in temperature and operating conditions.

L-oscillator with a resistor in the emitter circuit. This configuration (Fig. 8.12*a*) markedly reduces the dependence of pulse duration on the value of β.

To determine the pulse duration τ, we shall use the equivalent circuit in Fig. 8.12*b* where the dashed line encloses an ideal transformer which transforms the voltages across and the currents in the collector and base windings by a factor of $n = w_B/w_C$. The value of R_E is chosen sufficiently high ($R_E \gg r_{out(sat)}$ and $r_{in(sat)}$), so that $r_{out(sat)}$ and $r_{in(sat)}$ may be deemed short-circuited in our further discussion.

Referring to Fig. 8.12*b*, the collector and base currents are found to be

$$\begin{aligned} i_C(t) &= (1-n)^{-2} E_C (t/L_\mu + 1/R'_L + n^2/R_E) \\ i_B(t) &= nE_C/(1+n) R_E - i_C(t) \end{aligned} \tag{8.24}$$

Accordingly, Eq. (8.12) takes the form

$$\begin{aligned} &\tau_L \left[\frac{nE_C}{(1+n) R_E} - i_C(\tau) \right] \\ &\approx \tau_B E_C (1+n)^{-2} (\tau/L_\mu + 1/R'_L + n^2/R_E) \end{aligned} \tag{8.25}$$

because we neglect the minority-carrier excess charge. From Eq. (8.25) it follows that

$$\begin{aligned} \tau &= L_\mu (n^2/R_E) \left[\frac{a\beta(1+n)}{n} - 1 \right] - L_\mu/R'_L \\ &\approx (L_\mu n/R_E)(1 - R_E/nR'_L) \end{aligned} \tag{8.26}$$

where it is taken that $a\beta \gg 1$. The waveforms of currents and voltages existing over the interval τ appear in Fig. 8.12*c*.

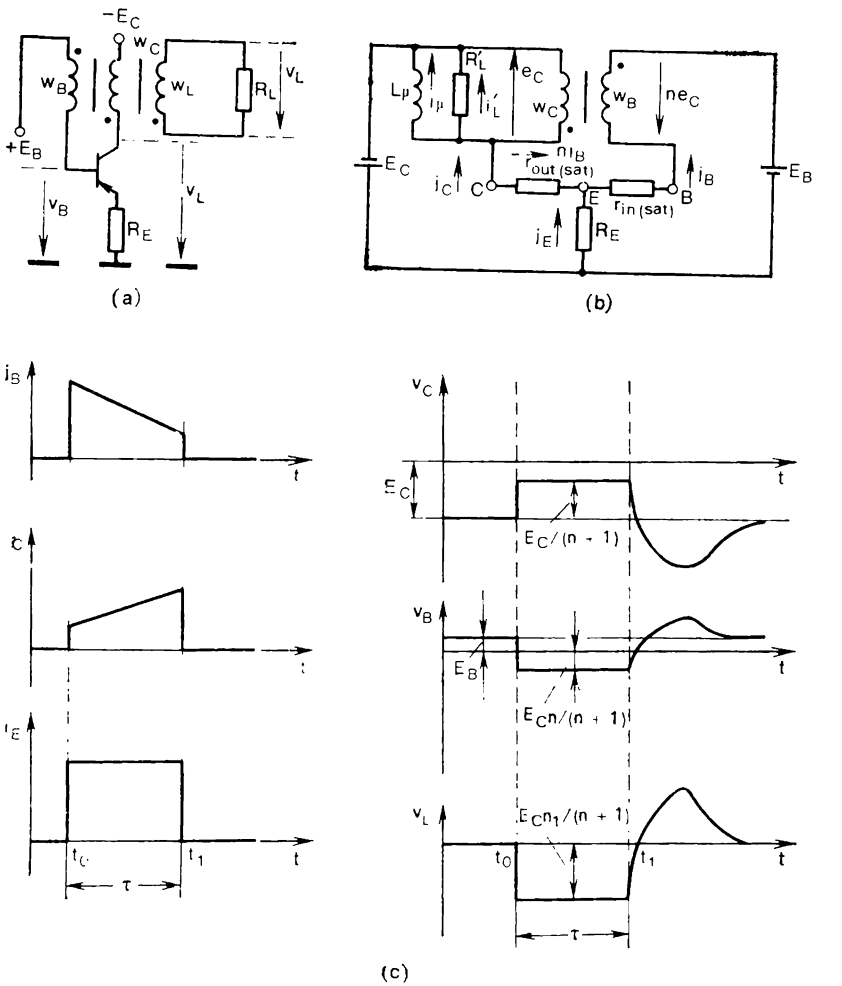


Fig. 8.12

The dependence of τ on load is minimized by inserting a buffer stage (Fig. 8.13). As the pulse top is shaped on transistor $T1$, the initial step change in collector current, $n^2E_C/(1+n)^2 R_E$, switches transistor $T2$ whose collector lead contains a variable load (say, a switchable number of ferrite-transistor cells) from cutoff to saturation, thereby permitting current to flow through the load.

The dependence of τ on L_μ and R_E can be minimized by using delay lines or saturable-core transformers. In the former case (Fig. 8.14), a short-circuited delay line with a sufficiently high

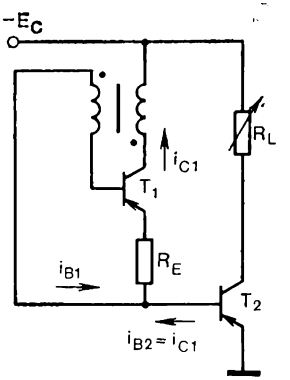


Fig. 8.13

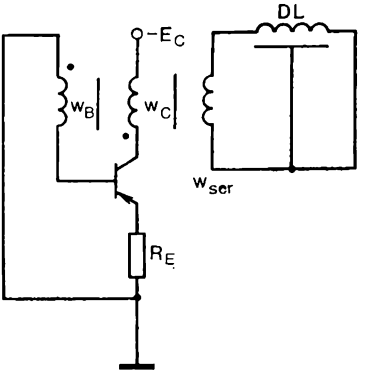


Fig. 8.14

characteristic impedance W is connected in the tertiary winding of the transformer. When the oscillator is excited, the line acts as a matched load, $R_L = W$, for the time interval required for the forward and backward wave to travel in the delay line, $2\tau_{dl}$. After the capacitances of the delay line discharge in the time $2\tau_{dl}$, the effective impedance of the delay line suddenly drops to the ohmic resistance of the inductances, thereby causing a rise in the collector current of the oscillator and the corresponding limiting charge, and the pulse top is terminated.

When use is made of an ordinary pulse transformer, the circuit parameters are chosen such that the required change in magnetic induction, ΔB (in the absence of saturation) will exceed $\Delta B_{max} = B_s - B_r$ for a given core:

$$\Delta B > \Delta B_{max} \tag{8.27a}$$

where $\Delta B = e_C \tau / (w_C S_{core})$ is the likely pulse duration in the absence of saturation. On substituting Eq. (8.26) for τ in Eq. (8.27a) and the expression for L_μ derived from Eq. (8.26), we may re-cast Eq. (8.27a) as

$$\Delta B_{max} < \frac{n}{1+n} (E_C/R_E) (1 - R_E/nR'_L) w_C \mu_\Delta / l_{core} \tag{8.27b}$$

When the condition defined by Eq. (8.27) is satisfied, the pulse duration τ [compare with Eq. (2.43)] is given by

$$\tau = \Delta B_{max} w_C S_{core}/e_C = (1 + n) \Delta B_{max} S_{core} w_C / E_C \tag{8.28}$$

At present, saturable-core pulse transformers are most commonly made with square-loop cores (Fig. 8.15*a*). Initially, when the transistor is OFF, the magnetization of the core is represented by point A $(-H_s, -B_s)$ in Fig. 8.15*b*, where H_s is due to $I_{bias} = E_C/R$

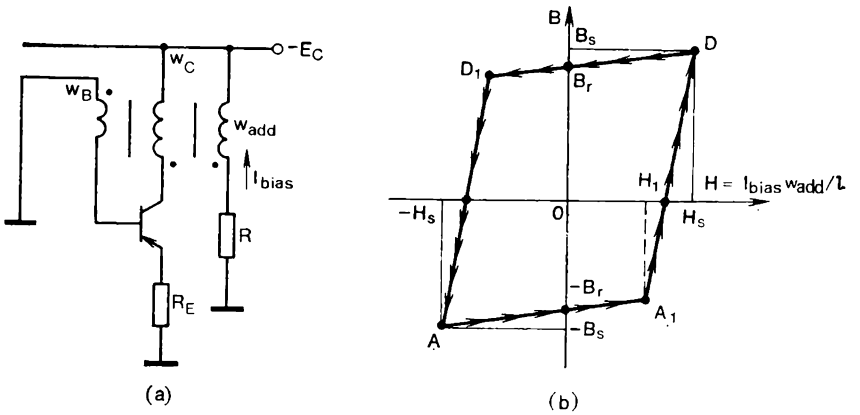


Fig. 8.15

traversing an additional winding, w_{add}. The trigger pulse switches the transistor into the active region, and the operating point begins to move along the lower portion of the hysteresis loop. The amplitude of the trigger pulse must be sufficiently large for the collector current to be such that the magnetizing field of the core reaches its threshold value, H_1 (point A_1).

Within the portion A_1D, the core has a high permeability, so the leading edge is formed in a cumulative manner as in an ordinary blocking oscillator. After the transistor has switched to saturation, the pulse top is shaped so that the representative point moves along the portion A_1D of the hysteresis loop as the magnetizing current increases. At point D, the core is saturated, the collector current suddenly rises, the transistor pulls out of saturation, and the pulse top is terminated. Owing to the fixed magnetic bias, the representative point moves along the portion DD_1 of the hysteresis loop, whereas the trailing edge is shaped within the portion D_1A and, after the transistor is driven to cutoff, the circuit recovers its original state.

Approximately, the pulse duration τ can be found from Eq. (8.28) in which ΔB_{max} must be replaced with $2B_s$:

$$\tau \approx (1 + n)\, 2B_s S_{core} w_C / E_C \qquad (8.29)$$

According to [3], the configuration shown in Fig. 8.15 may be used in ferrite-transistor gates. In [3] it is also pointed out that the increase in collector current as the pulse top is shaped is small, such a blocking oscillator may be treated as a current pulse generator. In fact, it is used as a clock generator for magnetic switches.

L-oscillators are frequently employed as square-wave generators or low-to-high d.c. voltage converters. Figure 8.16 shows the cir-

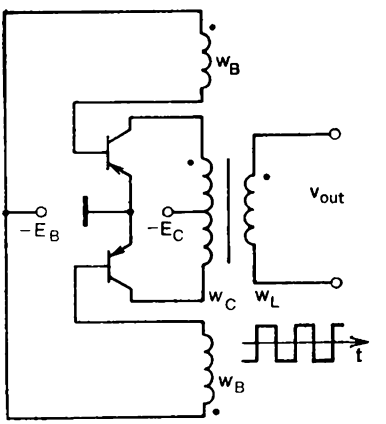


Fig. 8.16

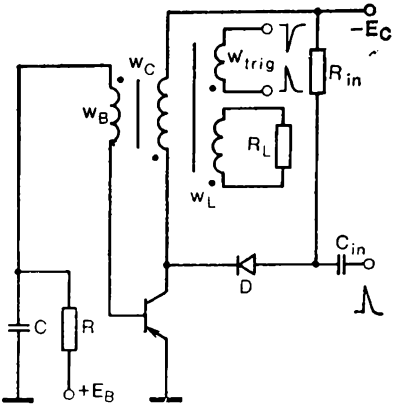


Fig. 8.17

cuit of a square-wave generator where two L-oscillators are interconnected into a push-pull circuit; one of them is triggered by the voltage backswing on the collector winding of the other. To convert a low collector voltage, E_C, to a higher value, the transformer uses a step-up load winding, and the voltage appearing across its output is rectified [4, 14].

Other modifications of the blocking oscillator are those using transformer coupling from collector to emitter or from emitter to base. Although they generate pulses with longer rise and fall times, they may be superior to other types as regards temperature stability of oscillation period and pulse duration, pulse duty factor, etc. (for more detail, see [3] and [14]).

Triggering and synchronization of blocking oscillators. In triggering a one-shot multivibrator or applying synchronizing pulses to a free-running blocking oscillator, care must be exercised to isolate the trigger pulse generator from the blocking oscillator. One way is to feed trigger pulses via a diode (Fig. 8.17) which will turn off owing to a rise in its cathode potential as the blocking oscillator shapes the leading edge of the output pulse. As an alternative, trigger pulses may be fed via a separate winding, w_{trig}, on the pulse transformer (see Fig. 8.17).

Vacuum-tube blocking oscillators. A vacuum-triode blocking oscillator similar in configuration to the circuit shown in Fig. 8.1*a* appears in Fig. 8.18. For one-shot operation, the polarity of E_g is reversed so that the tube is driven to cutoff in the absence of a trigger pulse. In general, a vacuum-tube blocking oscillator operates similarly to those built around transistors, except, of course, the aspects peculiar to vacuum tubes only. The tube of a blocking oscillator

emits in bursts, so that the pulse current may be a few tenths of an ampere to several amperes without exceeding the average power dissipation rating, if the pulse duty factor is small. The pulse top is usually shaped with the tube operating in the critical region, when the grid bias exceeds plate voltage and runs into tens or even hundreds of volts.

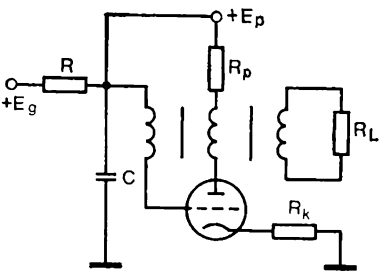


Fig. 8.18

For microsecond pulses, the tube may be regarded free from transit time effects, because the interelectrode capacitances are lumped with the stray capacitances of the transformer windings which determine the rise and fall times of the output pulses.

REVIEW QUESTIONS AND PROBLEMS

8.1. Draw the circuit of a collector-coupled blocking oscillator, explain its operation, and write the condition for its normal operation.

8.2. Write the basic design equations for a collector-coupled blocking oscillator and draw the applicable equivalent circuit.

8.3. Draw an equivalent circuit for a collector-coupled blocking oscillator to explain the shaping of the leading edge. Using this equivalent circuit, determine the rise time of the output pulse, and the optimal value, n_{opt}, of the transformation ratio corresponding to the minimum rise time.

8.4. Draw an equivalent circuit for a collector-coupled blocking oscillator to explain the shaping of the pulse top. Using this circuit, derive expressions for $Q(t)$ and $Q_{lim}(t)$. Using these expressions, determine the pulse duration for blocking oscillators generating short and long pulses, and also for an L-oscillator.

8.5. How is the trailing edge of a pulse shaped in a collector-coupled blocking oscillator?

8.6. Name the processes taking place between pulses in a collector-coupled blocking oscillator and determine the period of oscillation.

8.7. Draw the circuit of an L-oscillator with a resistor in the emitter lead and name its advantages. Derive the equation for the pulse duration and explain the advantages offered by this oscillator.

8.8. Draw the circuit and explain the operating principle of a blocking oscillator with an improved stability of pulse duration.

8.9. Draw the circuit and explain the operating principle of an L-oscillator used as a square wave generator.

8.10. Name triggering methods for blocking oscillators.

9 RAMP GENERATORS

9.1. GENERAL

A *ramp generator* is one that provides an output waveform a portion of which shows a linear variation of voltage or current with time (Fig. 9.1).

Ramp generators can be *triggered*, *free-running* or *synchronized*. A triggered ramp generator has an initial stable state to which it returns after a cycle of oscillation terminates. In this form of ramp generator, the *pulse repetition rate* or *frequency* is controlled by that of trigger pulses and may be constant (Fig. 9.1*a*) or variable (Fig. 9.1*b*). A free-running ramp generator switches back and forth between two levels, and its period of oscillation (Fig. 9.1*c* and *d*) is determined by the circuit parameters. A synchronized ramp generator, as its name implies, is timed by a clock control (or sync) voltage.

Sometimes, the direction in which the linear portion (*working stroke*) of the ramp waveform is varying is of importance. If this portion is positive-going, we have a linearly rising ramp (Fig. 9.1*a* and *c*); if it is negative-going, we have a linearly falling ramp (Fig. 9.1*b* and *d*). Frequently, the ramp voltage is called a *sawtooth voltage*. A very important application of such a waveform is in CRT oscilloscopes, where it serves to sweep the electron beam horizontally across the screen. Quite appropriately, it is called a *sweep voltage*. Because it varies linearly with time, another name for it is a *timebase voltage*.

The ramp (Fig. 9.2) can be specified in terms of the *sweep time* (the duration of the working stroke) τ, the initial level V_0, the maximum amplitude V_m, the rate of change (slope) of the sweep portion, and the *return* or *flyback time* that is the time τ_{fb} required for the return to the initial value prior to the start of the next cycle of oscillation. In practice, the ramp voltage will usually have a sweep

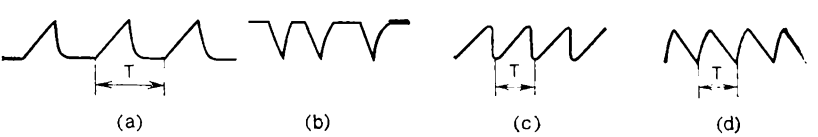


Fig. 9.1

time of a few tenths of a microsecond to tens of seconds, V_m from units to thousands of volts, a slope of 10^{12} to 10^{-2} V s^{-1}, and a flyback time of 1 to 50% of the sweep time.

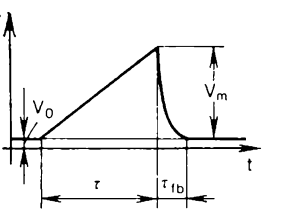


Fig. 9.2

Performance of Ramp Generators

Efficiency. A ramp generator should preferably generate an output voltage whose amplitude is large enough to make any subsequent amplification unnecessary as this usually distorts the output waveform. Accordingly, the efficiency of a ramp generator can be stated in terms of supply voltage utilization

$$K_E = V_m/E \tag{9.1}$$

The best performance in this respect is shown by transistor ramp generators for which K_E may exceed 99%.

Linearity. For most applications involving time measurement, it is essential that the rate of change of the ramp voltage should remain constant. Accordingly, the performance of a ramp generator is assessed in terms of the linearity of the ramp voltage (actually, its working stroke) and its stability with time and under varying operating conditions.

The ramp voltage $v(t)$ is said to be linear, if its rate of change (slope) during the working stroke remains constant. The deviation from linearity is expressed in terms of the slope or sweep-speed error

$$\xi = \frac{|\dot{v}(0) - \dot{v}(t)|}{\dot{v}(0)} = 1 - \frac{|\dot{v}(t)|}{|\dot{v}(0)|} \tag{9.2}$$

where $\dot{v}(0)$ and $\dot{v}(t)$ are the initial slope and the slope at time t (during the working stroke), respectively. Typically, ξ_{max} ranges between 0.1% and 10%, depending on the application of the ramp generator. Since $\xi_{max} \ll 1$, variations in the output voltage of a ramp generator during the working stroke may with sufficient acuracy be represented by the first three terms of a power series

$$v \approx A_1 t + A_2 t^2 + A_3 t^3 \tag{9.3}$$

Then, in view of Eq. (9.2),

$$\xi \approx t\,(2A_2 + \tfrac{3}{4}A_3 t)/A_1 \tag{9.4}$$

Variations in the rate of change (slope) is an absolute criterion of linearity. In time interval measurements, the nonlinearity of the ramp might lead to an error because the scale of the indicator is linearly calibrated and also because one has to interpolate between

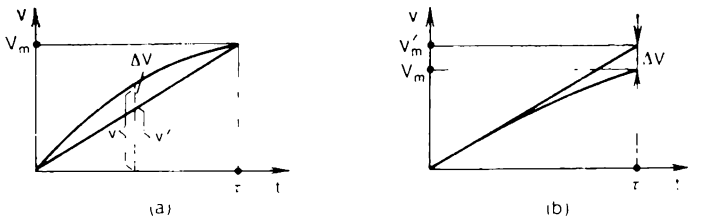


Fig. 9.3

adjacent calibration time marks. Therefore, the deviation from linearity may alternatively be characterized in terms of offset *error* δ_{oe} which shows how much the ramp waveform deviates from a straight line.

In contrast to the slope or sweep-speed error, the offset error depends not only on variations in the rate of change of the ramp voltage, but also on the relative position of the actual waveform and a reference straight line. Usually, the reference straight line is drawn so that it coincides with the actual waveform at the start and finish of the working stroke (Fig. 9.3*a*) or so that the actual and ideal ramp voltages have the same initial rate of change (Fig. 9.3*b*), that is,

$$\text{per cent } \delta_{oe} = \frac{|v(t) - v'(t)|}{V_m} \times 100 \text{ for Fig. 9.3}a \tag{9.5a}$$

$$\text{per cent } \delta_{oe} = \frac{|V'_m - V_m|}{V_m} \times 100 \text{ for Fig. 9.3}b \tag{9.5b}$$

where the primes mark the ideal ramp voltages.

Let us determine the maximum offset error when the working stroke is described by Eq. (9.3) and the ideal ramp voltage v' by the relation

$$v' = At \tag{9.6}$$

If the ideal and actual waveforms meet at the start and end (see Fig. 9.3*a*), we should set

$$\Delta V(\tau) = v(\tau) - v'(\tau) = 0$$

On finding the rate of change $A = A_1 + A_2\tau + A_3\tau^2$ of the ideal ramp voltage, we obtain the following relation between the absolute deviation and time:

$$\Delta V(t) = A_2t^2 + A_3t^3 - (A_2 + A_3\tau)\tau t \tag{9.7}$$

Differentiating Eq. (9.7) and equating the derivative to zero, we obtain the instant at which the error is a maximum

$$t_0 = -\frac{A_2}{3A_3} + \sqrt{\left(\frac{A_2}{3A_3}\right)^2 + \frac{A_2\tau + A_3\tau^2}{3A_3}} \tag{9.8}$$

If the second term under the radical sign in Eq. (9.8) is smaller than the first, then

$$t_0 \approx 0.5\tau\,(1 + A_3\tau/A_2) \tag{9.9}$$

When the terms of the polynomial (9.3) decrease monotonically the slope error may be determined with an accuracy sufficient for practical purposes by assuming that $A_3 = 0$. Then, at time $t_0 \approx 0.5\tau$, the maximum offset error is

$$\delta_{oe,\,max} \approx A_2\tau/4A_1 \tag{9.10}$$

Because when $A_3 = 0$ the slope error is a maximum at the end of the working stroke, the maximum values of the slope error and the offset error are connected by a relation of the form

$$\xi_{max} \approx 8\delta_{oe,\,max} \tag{9.11}$$

When the ideal and actual waveforms meet at one point (Fig. 9.3*b*), $A = A_1$, the maximum offset error is

$$\delta_{max} = \tau|A_2 + A_3\tau|/A_1 \tag{9.12}$$

and

$$\xi_{max} \approx 2\delta_{oe,\,max} \tag{9.13}$$

Thus, in either case the slope error is a more rigorous test than the offset error. In our further discussion, we shall use the maximum values of the slope and offset error omitting the subscript "max" for brevity.

Stability of the ramp waveform. Assuming a high degree of linearity, variations in the ramp waveform can be stated in terms of the relative variation in the initial rate of change, $\delta\xi$, and in the initial voltage, δV_0:

$$\delta\xi = \frac{|\dot{v}(0) - \dot{v}'(0)|}{\dot{v}(0)}$$
$$\delta V_0 = \frac{|V_0 - V_0'|}{V_0} \tag{9.14}$$

where $\dot{v}'(0)$ and $V'(0)$ are the rate of change of the ramp and the initial voltage modified to accommodate the instability of circuit parameters.

As is seen, $\delta\xi$ and δV_0 characterize the stability of the working strokes of the ramp waveform.

Other characteristics of ramp generators. These include speed, loading capacity, versatility, and economy. Because the sweep time, τ, is always specified in advance, the speed of a triggered ramp generator is characterized in terms of the pulse duty factor: $\eta = \tau/T_{min}$, where T_{min} is the minimum permissible repetition period of trigger pulses. In complex ramp generators, in addition to the flyback time,

the factors preventing an increase in the pulse duty factor include the recovery time τ_{re} required for any transients in the circuit to die out and for its voltages and currents to reach their steady-state values. For free-running ramp generators the speed is usually the frequency of free oscillations.

The loading capacity refers to the effect that load has on the accuracy of a ramp generator. Versatility refers to provision and convenience of speed adjustment over a wide range. The economy of a ramp generator refers to the count of circuit components used and the power consumed by the generator.

Methods of Ramp Generation

At present, practically all methods of ramp generation are based on letting a capacitor charge or discharge at a constant (or nearly constant) current. Variations in the voltage $v_\frown$ across a capacitor C as compared with its initial value V_0 is proportional to the charge stored:

$$v_\frown = \frac{1}{C}\int_0^t i\,dt$$

If the current charging the capacitor is held constant, $i = I_0$, the capacitor voltage will change linearly, $v_\frown = At$, where

$$A = I_0/C \tag{9.15}$$

To obtain a train of ramp pulses, the charge and discharge circuits of the capacitor must be switched at the start ($t = 0$) and the end ($t = \tau$) of the stroke, or sweep. Thus, functionally, the circuit of any ramp generator (Fig. 9.4) must contain two basic elements: a source of constant charging (discharge) current and a switch.

A triggered ramp generator (Fig. 9.4*a*) can operate only when it is driven by either rectangular pulses of duration τ, which disable the switch for the duration of the sweep time, or short trigger pulses. In the latter case, a ramp generator should contain a positive feedback path (shown by the dashed line in Fig. 9.4*a*) which opens the switch for the duration of the sweep time. In either case, the input voltage causes the ramp to start at precisely the same time ($t = 0$), which is essential in precise time measurements.

Positive feedback is likewise used in a free-running sawtooth voltage generator (Fig. 9.4*b*). The generator may be synchronized either at the fundamental frequency of the synchronizing voltage or at one of its submultiples. As a rule, the synchronizing voltage in sawtooth generators controls the start of the return stroke ($t = \tau$), so they are not suited for precise time measurement and mostly serve as timebase generators in TV and oscilloscopes.

Stabilization of the charging current. According to the manner in which the capacitor charging (or discharge) current is stabilized,

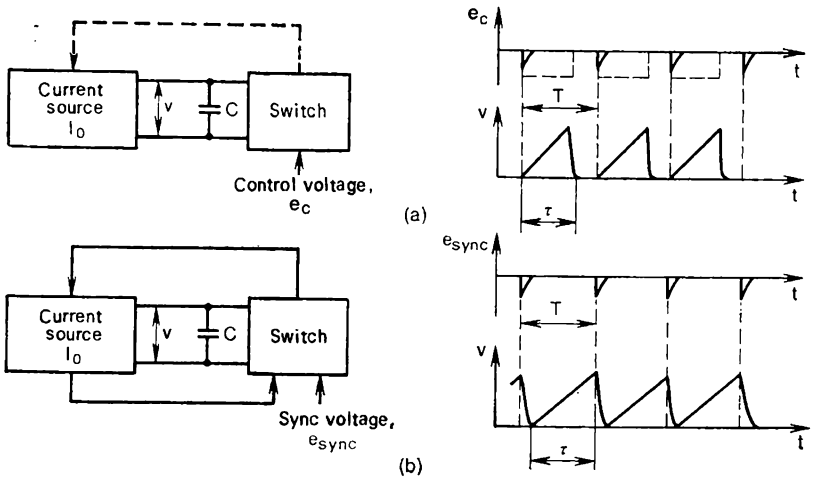


Fig. 9.4

it is convenient to divide all existing ramp generators into three general groups. The first group includes ramp generators containing a simple integrating *CR*-network, in which the capacitor charges (discharges) through a relatively high-value series resistor from a constant current source. These *CR* ramp generators are the simplest and, at the same time, least accurate of all ramp generators. The second group covers ramp generators in which the discharging (or charging) resistor is replaced by a constant-current circuit and the capacitor charges via a nonlinear element having a high incremental (differential) resistance. Constant-current ramp generators fall in the medium-accuracy class. The third group covers ramp generators which utilize positive and negative feedback amplifiers to supply a compensating voltage to keep the charging current constant to a high degree of accuracy (usually called *bootstrap ramp generators*).

9.2. *CR* RAMP GENERATORS

Consider the simplest ramp generator (Fig. 9.5) in which opening the switch *Sw* allows capacitor *C* to charge via series resistor *R* from a d.c. voltage supply *E*. Assume that when the switch is closed, the initial voltage on the capacitor is $v = V_0$.

Working stroke (sweep portion) of the ramp voltage. As the switch opens ($t = 0$), the capacitor begins to charge exponentially

$$v = (E - V_0)\,[1 - \exp(-t/\tau_0)] + V_0 \qquad (9.16)$$

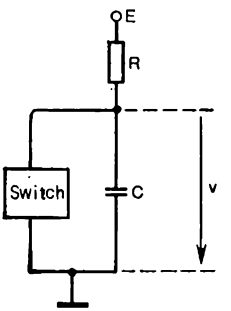


Fig. 9.5

where $\tau_0 = RC$. Obviously, in this circuit the voltage varies linearly within the relatively small initial portion of the exponential curve defined by Eq. (9.16) for which

$$\tau/\tau_0 \ll 1 \tag{9.17}$$

Retaining only the quadratic term in the convergent power series of the exponential function in Eq. (9.16), we obtain

$$\begin{aligned} A_1 &= (E - V_0)/\tau_0 \\ A_2 &= (E - V_0)/2\tau_0^2 \end{aligned} \tag{9.18}$$

Thus, if the circuit of Fig. 9.5 satisfies the condition defined by Eq. (9.17), opening the switch will cause a rise in the voltage for which the initial rate of change, slope error and utilization factor are given by

$$\begin{aligned} \dot{v}(0) &= A_1 = (E - V_0)/\tau_0 \\ \xi &\approx \tau/\tau_0 \\ K_E &\approx (1 - V_0/E)\,\xi \approx \xi \quad (\text{for } V_0 \ll E) \end{aligned} \tag{9.19}$$

Effect of loading. In a practical CR ramp generator, the charging capacitor is connected in parallel with the resistance of the switch (when it is open) and the load resistance. The total impedance shunting the charging capacitor may be depicted by two parallel arms respectively containing the resultant stray capacitance C_s and the load resistance R_L (Fig. 9.6*a*). Lumped with the capacitance of the charging capacitor, the stray capacitance affects the initial rate of change of the ramp voltage:

$$\dot{v}(0) = \frac{E}{R(C + C_s)} \approx (I_0/C)(1 - C_s/C) \tag{9.20}$$

where $C_s/C \ll 1$.

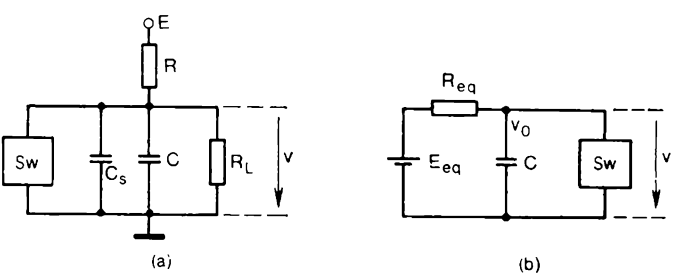


Fig. 9.6

The stray capacitance also includes variable components, such as the interelectrode capacitances of the electron devices. Because of this, the rate of change of the ramp voltage can further be stabilized by choosing C to be several times the value of C_s.

To evaluate the effect of the ohmic component of the load resistance on the circuit of Fig. 9.6*a*, we shall use Thevenin's theorem to obtain the circuit of Fig. 9.6*b* where

$$R_{eq} = R \parallel R_L \text{ and } E_{eq} = ER_{eq}/R \tag{9.21}$$

Hence,

$$\begin{aligned} \dot{v}(0) &= E/R - V_0/R_{eq} \\ &= I_0 \left(1 - \frac{V_0}{E} \frac{R_L}{R_L + R}\right) \end{aligned} \tag{9.22}$$

It follows from Eq. (9.22) that the ohmic component of the load resistance affects the initial rate of change of the ramp voltage only in the presence of the initial voltage V_0. Because of this, every effort must be exercised to keep V_0 to a minimum. Basically, R_L leads to an increase in the slope error to $\xi_{eq} = \tau/\tau_{eq}$, where $\tau_{eq} = R_{eq}C$, or in relative terms

$$\xi_{eq}/\xi = 1 + R/R_L \tag{9.23}$$

As follows from the foregoing, a major disadvantage of a *CR* ramp generator is the low utilization of supply voltage ($K_E \approx \xi \ll 1$), although the linearity is satisfactory, and a low loading capacity. In other words, this form of ramp generator produces a sufficiently linear ramp only at $R_L \gg R$.

Return stroke of the ramp. An important characteristic of the return stroke is the return or flyback time, τ_{fb} as it determines the speed of the ramp generator. In ordinary low-voltage ramp generators, the switches usually are diodes and transistors connected in a common-emitter circuit (see Chap. 3). In a ramp generator using a diode switch (Fig. 9.7*a*), the initial voltage on the capacitor at time $t < t_0$ may be deemed negligibly small if the bias voltage E_{bias} is chosen such that $E_{bias} = ER_{in}/R$, where R_{in} is the blocking resistor. The working stroke of the ramp occupies the interval from t_0 to t_1, when the diode is driven to cutoff by the input voltage pulse of amplitude $E_{in} \geqslant V_m$, applied to the base. The return stroke begins at time t_1, when the diode turns on after the input pulse ceases. If the blocking capacitor C_{in} is sufficiently large [$C_{in}(R_s + R_{in})$ exceeds both τ and τ_{fb}], it may approximately be deemed that the capacitor discharges exponentially with a time constant $\tau_{re} = CR_{re}$, where $R_{re} = R \parallel R_{in} \parallel R_s$, and the flyback time will be

$$\tau_{fb} = k_{re}\tau_{re} \tag{9.24}$$

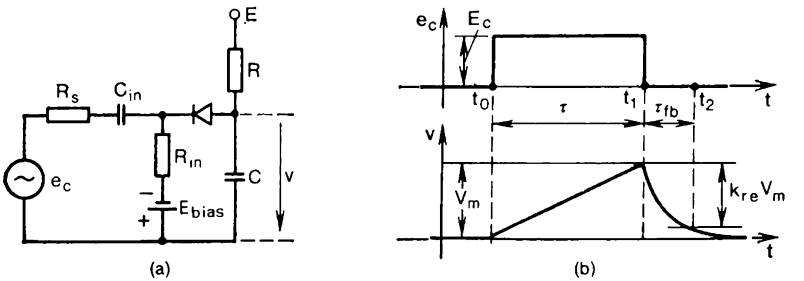


Fig. 9.7

where

$$k_{re} = -\ln (1 - k_{\text{dec}})$$

and

$$k_{\text{dec}} = \frac{V_m - v(t_2)}{V_m} = v_\frown(t_2)/V_m$$

As already noted in Chap. 2, if the required decay level is $k_{\text{dec}} = 0.95$, then k_{re} may be taken as 3.

It is useful to relate the sweep and flyback times of the ramp. From Eqs. (9.19) and (9.24) we find that

$$\tau_{fb} = \tau k_{re} R_{re}/RK_E \tag{9.25}$$

It follows from Eq. (9.25) that, given τ_0, the flyback time can be reduced by increasing the value of R and decreasing the value of C.

In a ramp generator using a transistor switch (Fig. 9.8), the transistor is initially at saturation and $v = V_{\text{CE(sat)}} \approx 0$. The positive-going change of input voltage E_{in} applied at time t_0 turns off the transistor after a time τ_{OFF} which is the sum of τ_s and τ_t where τ_s is the storage time (the time required for removal of the minority-carrier excess charge) and τ_t is the transition time of the transistor.

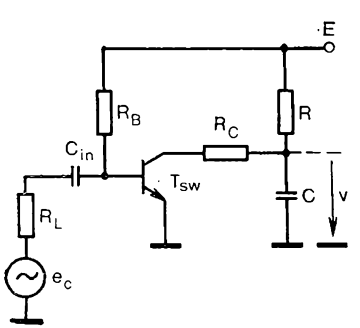


Fig. 9.8

When the collector current of the transistor ceases, the voltage begins to rise exponentially at the initial rate $v(0) \approx E/\tau_0$.

During the return stroke when the capacitor discharges, transistor T_{sw} may be saturated or active, depending on the value of the collector resistor R_C (Fig. 9.9). Initially, at $t < t_0$, the switch is saturated, and for the load $R + R_C$ the operating point lies at the intersection, O, of the load line with the saturation curve. If R_C

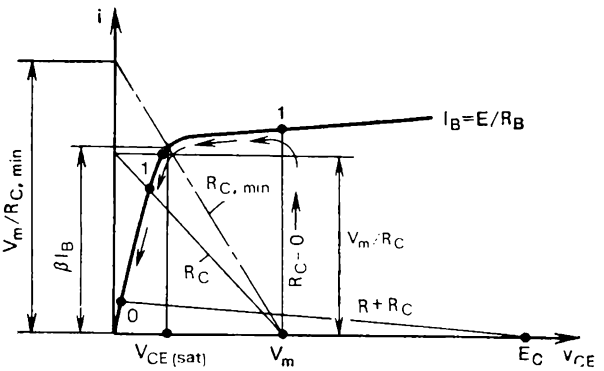


Fig. 9.9

$\geqslant R_{C,min}$, the position taken up by the operating point immediately at the time $(t_1 + 0)$, that is, the end of the working stroke, will be the intersection of the line V_m/R_C with the *V-I* characteristic of the transistor for $I_B = E/R_B$ (point *1*) and will not go beyond the saturation region. As the capacitor discharges, the operating point moves down the saturation line towards the initial point 0. In this case, sometimes termed the saturated switch case, the flyback time will be defined by Eqs. (9.24) and (9.25) for a ramp generator with a diode switch, where $R_{re} = (R_C + r_{out(sat)}) \parallel R \approx R_C$. Noting that $r_{out(sat)} \ll R_C$, we can determine the minimum value of the collector resistor, $R_{C,min}$, in the saturated switch. Since it is required that $V_m/R_{C,min} = \beta I_B$ and referring to Fig. 9.9, it is found to be

$$R_{C,\,min} = V_m/\beta I_B = K_E R_B/\beta \tag{9.26}$$

When $R_C = 0$, the operating point occurs at point *1* within the active portion of the *V/I* characteristic of the transistor at the beginning of the return stroke, then moves to the left along this characteristic as far as the point of inflection, and goes back to its initial position. Noting that $V_{C(sat)} \ll V_m$, we may deem that in this case (which may be called the unsaturated switch case) practically all of the return stroke is completed while the transistor is in the active region. The equivalent circuit that can be used to determine the flyback time for a ramp generator using an unsaturated switch is shown in Fig. 9.10. From inspection of the equivalent circuit, we may write the following exponential relation for a recharge of the capacitor:

$$v = -(\beta I_B - I_0)\, R_{sh} + [(\beta I_B - I_0)\, R_{sh} + V_M] \exp\,[-(t - t_1)/\tau_{re}] \tag{9.27}$$

where $I_0 = E/R$, $R_{sh} = r_{out} \parallel R$, and $\tau_{re} = CR_{sh}$.

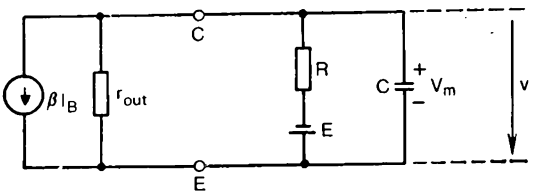


Fig. 9.10

On substituting $v\,(t_2) = v\,(\tau_{fb}) \approx 0$ in Eq. (9.27), we get

$$\tau_{fb} = \tau_{re} \ln [1 - V_m/R_{sh}\,(\beta I_B - I_0)] \qquad (9.28a)$$

Because the second term in the brackets of the logarithm is small in comparison with unity, we may write

$$\tau_{fb} \approx \tau/\,(K_{sat} - 1) \qquad (9.28b)$$

where, as with a diode switch, we may take it that $\tau \approx CV_m/I_0$.

9.3. CONSTANT-CURRENT RAMP GENERATOR

The major disadvantage of the simple CR ramp generator is its nonlinearity. To produce a linear ramp, the capacitor charging current must be held constant. This can be achieved by replacing the charging resistor with a *constant current circuit.*

One possible source of constant current is a d.c.-stabilized common-emitter transistor circuit such as shown in Fig. 9.11*a*. The role of the charging resistance is played by the resistance of the collector-base junction, and the initial rate of change of the capacitor charge is determined by the initial collector current I_0.

Consider the operation of the constant-current ramp generator. To maintain I_0 constant, the emitter lead contains a voltage supply $E \gg e_0$ and an emitter resistor $R \gg r_{in}$. In the absence of a trigger pulse, the switching transistor $T2$ is in the active region and acts as an emitter follower operating into a high-value capacitive load.

Initially, at $t < t_0$ (Fig. 9.11*b*), the current flowing in the collector leads of the ON transistors $T1$ and $T2$ is $I_0 \approx E/R$, so the initial capacitor voltage is close to E_C [$V_0 \approx E_C - \alpha_1 E R_B/R\,(1 + \beta_2)$]. At time t_0, transistor $T2$ is turned off by the leading edge of the trigger pulse, and the capacitor begins to discharge via the collector-base circuit of $T1$ and the shunting resistances, namely $r_{out,c\text{-}o}$ of $T2$ and the load resistance R_L. If the duration τ of the trigger pulse does not exceed the maximum duration of the working stroke ($\tau_{max} \approx CE_C/I_0$), the amplitude of the ramp will be lower than the max-

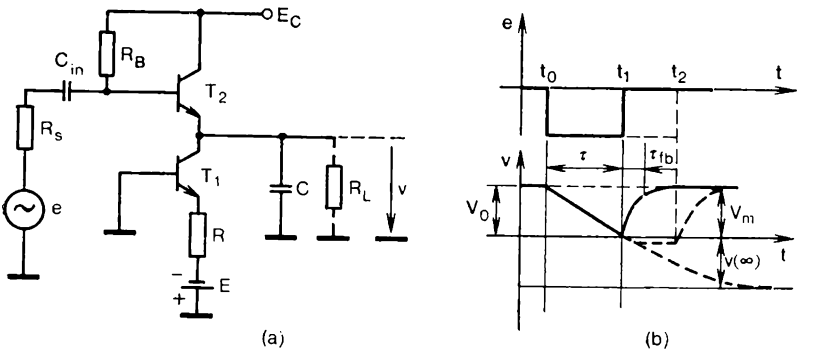


Fig. 9.11

imum one ($V_m < V_{m,max}$), and the return stroke, that is, the charging of the capacitor via the emitter circuit of $T2$, will begin at time t_1 (when the trigger pulse ceases), immediately at the end of the working stroke. If $\tau > \tau_{max}$, the end of the working stroke and the start of the flyback will be separated by an interval during which the output voltage will be $v = -V_{CB(sat)}$, where $V_{CB(sat)}$ is the saturation voltage of $T1$ in the absence of a voltage supply between collector and base.

Working stroke of the ramp. To determine the working-stroke characteristics of the ramp, we shall refer to Fig. 9.12 where the input and output circuits of transistor $T1$ are represented by quasi-static equivalent circuits and the shunt resistor is $R_{sh} = r_{out,c-o} \parallel R_L$. From analysis of the circuit in Fig. 9.12 it follows that the re-charging of the capacitor follows a linearized exponential curve with fairly high values of the limiting voltage and time constant

$$v(t) = v(\infty) - [v(\infty) - v(0)] \exp[-(t - t_0)/\tau_{eq}] \qquad (9.29)$$

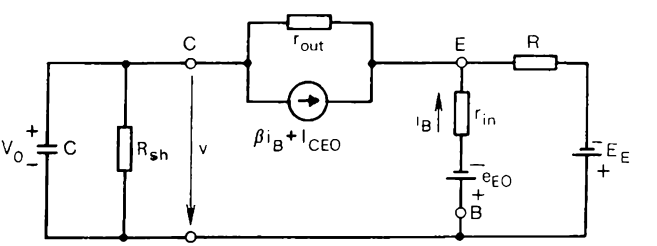


Fig. 9.12

where

$$\tau_{eq} \approx CR_{eq}$$

$$v(0) = V_0$$

$$v(\infty) \approx -\frac{(1+\beta) R_{sh} r_{out} [I_{CO}R + \alpha (E - e_{EO})]}{R [R_{sh} + (1+\beta) r_{out}]}$$

$$r_{in} \ll R,\ R_{sh},\ \beta r_{out}$$

$$R \ll r_{out}$$

$$R_{eq} = R_{sh} \parallel [(1 + \beta) r_{out}]$$

It follows from Eq. (9.29) that the initial rate of change of the ramp voltage during the working stroke is

$$\dot{v}(0) = I_0/C \tag{9.30}$$

where

$$I_0 = -\frac{1}{C}\left(V_{EO}/R_{sh} + I_{CO} + \alpha \frac{E - e_{EO}}{R}\right)$$

For the initial rate of change of the ramp to remain sufficiently stable, it is required that the third term in the brackets of Eq. (9.30) should be prevailing:

$$\alpha \frac{E - e_{EO}}{R} \gg I_{CO} + V_0/R_{sh} \tag{9.31}$$

Then, since $\alpha \approx 1$ and $E \gg e_{EO}$, the initial collector current I_0 determining the initial rate of change of the ramp will to a first approximation be independent of the transistor parameters and load:

$$I_0 \approx E/R$$

Nonlinearity. In view of Eqs. (9.4) and (9.29), the nonlinearity (slope error) is given by

$$\xi = \tau/(CR_{eq}) \tag{9.32a}$$

On replacing the sweep time by an approximate relation, $\tau \approx V_m \tau_0/E$, and introducing K_E which for the circuit of Fig. 9.11 should preferably be defined as $K_E = V_m/E\,(1 + \lambda)$, where $\lambda = E_C/E$, we get

$$\xi = K_E (1 + \lambda) R/R_{eq} \tag{9.32b}$$

It follows from Eqs. (9.19) and (9.32b) that the circuit of Fig. 9.11 will be effective only if the ramp generator is loaded into a sufficiently high resistance: $R_{eq} \approx R_L > (1 + \beta) r_{out}$. The low loading capacity is a major limitation or the constant-current ramp generators in Fig. 9.11.

Parameter optimization. It follows from Eq. (9.32a) that in order to reduce ξ, the value of the charging capacitor should be increased and that of R decreased so that their product, $\tau_0 = CR$,

which controls the initial rate of change of the ramp, remains constant. However, an increase in C entails an increase in the flyback time. In the general case, the diode, saturated, and emitter-follower switches may be arranged so that

$$\tau_{fb} = k_{re} R_{sw} C \qquad (9.33)$$

where the value of R_{sw} depends on the type of switch used. For the circuit of Fig. 9.11, $R_{sw} \approx R_B/(1 + \beta)$. Thus, assuming that the maximum permissible flyback time is

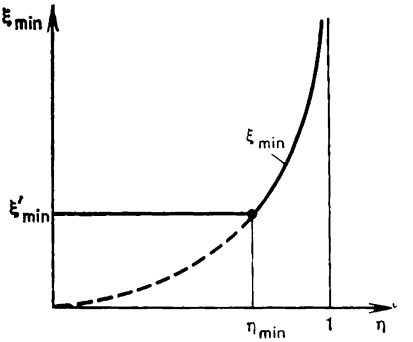


Fig. 9.13

$$\tau_{fb,\,max} = T - \tau_{max}$$

where T is the switching period of the ramp generator, the maximum capacitance of the charging capacitor with which the flyback time does not exceed $\tau_{fb,max}$ is

$$C_{max} = (T - \tau)/(k_{re} R_{sw}) \qquad (9.34)$$

Substituting $C = C_{max}$ in Eq. (9.32a) gives the minimum slope error as a function of the speed of the ramp generator (Fig. 9.13, Table 9.1):

$$\xi_{min} = \frac{\eta}{1-\eta} k_{re} R_{sw}/R_{eq} \qquad (9.35)$$

Table 9.1. Ramp Linearity as a Function of Speed for $k_{re} = 3$ and $R_{sh}/R_{eq} = 0.02$

η	$\eta/(1-\eta) = \tau/\tau_{fb}$	ξ_{min}, %
0.09	0.1	0.6
0.5	1	6
0.6	1.3	8
0.7	2.3	14
0.8	4	24
0.9	9	54

Another factor setting the limit for an increase in C and a proportionate decrease in R is the maximum safe collector current $I_{C,safe} \approx I_0$, of the constant-current transistor. Substituting

$$C'_{max} = I_{C,\,safe}/\dot{v}\,(0) \qquad (9.36)$$

in Eq. (9.32a) gives

$$\xi'_{min} = K_E (1 + \lambda)\, E/(R_{eq} I_{C,\,safe})$$

On the plot of Fig. 9.13, ξ'_{min} corresponds to the minimum pulse duty factor, η_{min}, just sufficient for obtaining a maximum linearity in accord with Eq. (9.35), and with C_{max} chosen by Eq. (9.34). Taking typical values, that is, $K_E = 1$, $\lambda = 2$, $R_{eq} = 10^4\ \Omega$, $E = 20$ V, and $I_{C,safe} = 50$ mA, we get $\xi'_{min} = 8\%$ and $\eta_{min} = 0.7\%$.

9.4. BOOTSTRAP AND MILLER INTEGRATOR RAMP GENERATORS

The initial charging current of the capacitor in a ramp generator can be maintained constant by inserting a source of compensating emf v_{comp} in series (Fig. 9.14*a*) or in parallel (Fig. 9.14*b*) with the charging (discharging) path of the capacitor. In the circuit of Fig. 9.14*a*, the compensating emf is supplied by an amplifier whose output voltage must be an exact copy of the voltage across the charging capacitor, $v_{comp} = v_C$. The input voltage v_{in1} is taken to be that across points a_1 and a_2. This configuration is not fully satisfactory, because the potential at either terminal of the charging voltage source E varies relative to the virtual ground as the ramp is generated. The charging current, $i_C = i$, will therefore remain constant, $i_C = I_{CO} = E/R$, only if the amplifier gain is unity, $K_g = 1$. Obviously, the effect of variations in the potential at point a_1 on I_C will be compensated for by proportional variation in the potential at a_3 owing to the positive feedback path which includes the supply E. In Soviet practice, this configuration using a noninverting amplifier is known as a *series bootstrap ramp generator.*

The arrangement of Fig. 9.14*b* likewise uses a noninverting amplifier for which $v_{in} = v_C$ and $v_{out} = K_g v_{in} = v_C$. It is an easy matter to see that for $i_C = I_{CO}$ to remain constant, it is essential that the amplifier gain should be $K_g = 1 + R_s/R$. In Soviet usage, the form of circuit is known as a *parallel bootstrap ramp generator.*

In practice, the value of K_g is chosen to be different from the above values deduced for ideal circuits, so as to make up for the effect of stray parameters.

If we return one side of the supply E (point a_3) to ground in the circuit of Fig. 9.14*a*, the input voltage for the amplifier will be

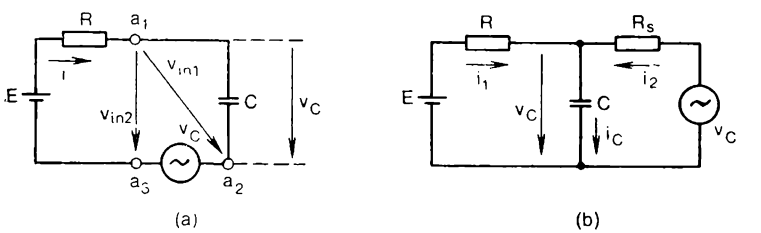


Fig. 9.14

that across points a_1 and a_3, that is, v_{in2}. It can readily be seen that the charging current of the capacitor will remain constant only if the potential at a_1 remains unchanged for the duration of the ramp time, and this can only be achieved if the amplifier gain is infinitely large. This requirement is met by utilizing the Miller effect, so the respective circuit is known as the *Miller integrator ramp generator*. In practice, the Miller integrator ramp generator often proves more accurate and effective than the bootstrap ramp generator.

Series Bootstrap Ramp Generator

A generalized circuit of series bootstrap ramp generators is shown in Fig. 9.15. Since, as already noted, it is difficult to isolate from ground both sides of the supply, the latter is replaced by the initial voltage across the bootstrapping capacitor C_1, that is, $V_{C1,0}$. The initial charge on this capacitor is restored at the end of flyback from the supply E via switch $Sw2$ and the output resistance R_{out} of the amplifier. Switch $Sw1$ controls the discharge circuit of the charging capacitor C.

Initially at $t < t_0$, the switches are closed and a constant current I_0 is flowing around the circuit containing the switches and the charging resistor. Now the initial voltage on the bootstrapping capacitor is close to E, and a relatively low initial potential V_0 is maintained on the charging capacitor C. At $t = t_0$, the switches open, and the voltage across the circuit composed of the charging RC network and the output circuit of the amplifier gives rise to an initial current which is equal, in a first approximation, to I_0. The rise in voltage on the charging capacitor entails a proportionate increase in the output potential, v_{out}, of the amplifier, which is transferred via C_1 to the upper terminal of the charging resistor R. If the potential at the upper terminal of R rises at the same rate as the

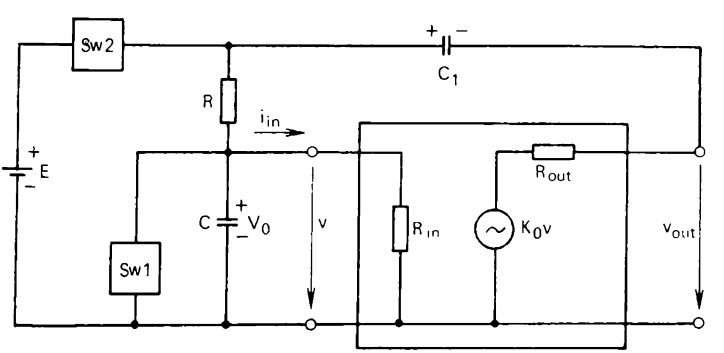


Fig. 9.15

voltage on C, the charging current flowing in the above mentioned circuit will remain constant, and the slope of v will likewise remain constant. Since the amplifier produces no limiting action, the ramp will be terminated by closure of $Sw1$ through which the capacitor discharges during flyback.

As a rule, $Sw2$ closes at the end of the return stroke. In the meantime, the nearly constant current I_0 causes the charge on C_1 to decrease by $\Delta Q \approx I_0 (\tau + \tau_{fb})$ during flyback in the same manner as it does during the ramp time. Upon closure of $Sw2$, this charge is restored, and the circuit returns to its original state.

Let us see what parameters the circuit should have in order to generate a ramp voltage, v. Reckoning time from the start of the working stroke and assuming that the switches are open, inspection of the circuit in Fig. 9.15 leads us to the following Laplace transform of the ramp voltage:

$$v(s) = V_{C1,0}/C(R + R_{out})(s - s_1)(s - s_2) \tag{9.37}$$

where s_1 and s_2 are the roots of the characteristic equation

$$s^2 + s\frac{K}{C(R + R_{out})} + \frac{1}{C_1 R_{in} C(R + R_{out})} = 0 \tag{9.38}$$

where $K = 1 - K_g + C/C_1 + (R + R_{out})/R_{in}$ is the linearization variable.

According to the manner in which the circuit of Fig. 9.14*a* operates, the linearization variable should be reduced and the values of C_1 and R_{in} should be increased. Then,

$$s_1 \approx -(KC_1R_{in})^{-1}$$

$$s_2 \approx -K/C(R + R_{out})$$

and the roots are such that $|s_1|^{-1} \ll |s_2|^{-1}$.

The sweep time τ will be such that

$$\tau \ll KC_1R_{in} \tag{9.39}$$

Neglecting $|s_1|^{-1}$ as being very small, we may approximately write the original time function for the Laplace transform defined by Eq. (9.37) as

$$v(t) \approx E_{eq}[1 - \exp(-t/\tau_{eq})] \tag{9.40}$$

where $\tau_{eq} = C(R + R_{out})/K$ and $E_{eq} = V_{C1,0}/K$.

It follows from Eq. (9.40) that positive voltage feedback causes the charging resistance and the charging voltage to increase by a factor of $1/K$.

The initial rate of change and the nonlinearity of the ramp can be found from Eq. (9.40):

$$\dot{v}(0) = I_0/C = \frac{V_{C1,0}}{\tau_0 (1 + R_{out}/R)} \tag{9.41a}$$

$$\xi = \tau/\tau_{eq} = K_E K \tag{9.41b}$$

where

$$I_0 = \frac{V_{C1,0}}{\tau_0 (1 + R_{out}/R)}$$

$$\tau_0 = RC$$

$$K_E = V_m/V_{C1,0}$$

Thus, it follows from Eq. (9.41) that the bootstrap ramp generator in Fig. 9.15 improves the linearity by a factor of $1/K$ in comparison with an CR ramp generator. From Eqs. (9.40) and (9.41) it is also seen that this circuit can generate a precisely linear ramp if K_g is chosen such that

$$K = 1 - K_g + C/C_1 + (R + R_{out})/R_{in} = 0 \tag{9.42}$$

The condition defined by Eq. (9.42) might be satisfied by including, say, a two-stage amplifier having the required gain. In practice, however, the nonlinearity and instability of the transistors and other circuit components would inevitably violate the condition (9.42). In fact, even a minute change in the parameters (say, K_g) would substantially affect the linearity of the ramp. To minimize this effect, K_g must be stabilized by applying a large amount of negative feedback.

In practice, the function of an amplifier with a gain of close to unity can be performed by an emitter follower or an operational amplifier.

The circuit of a series bootstrap ramp generator using an IC opamp is shown in Fig. 9.16*a*. The diode D serves as a second switch held open for the duration of the ramp and the larger proportion of the flyback (the interval from t_0 to t_2 in Fig. 9.16*b*) because its cathode is pulled up in potential above its anode as the ramp is passed through the opamp. The main switch is an unsaturated switching transistor T (see Sec. 9.2).

At time t_0, the leading edge of the input pulse turns off the transistor switch, and the potential at the noninverting input of the opamp begins to rise at an initial rate controlled by the charging current previously flowing through the closed switch. Because the opamp output is coupled to the diode cathode via a high-value capacitor C_1, the gain of the opamp is low at first, and remains so until the diode turns off, and the voltage V varies at a considerably lower rate than v_{in}. In the meantime, the feedback path is practically

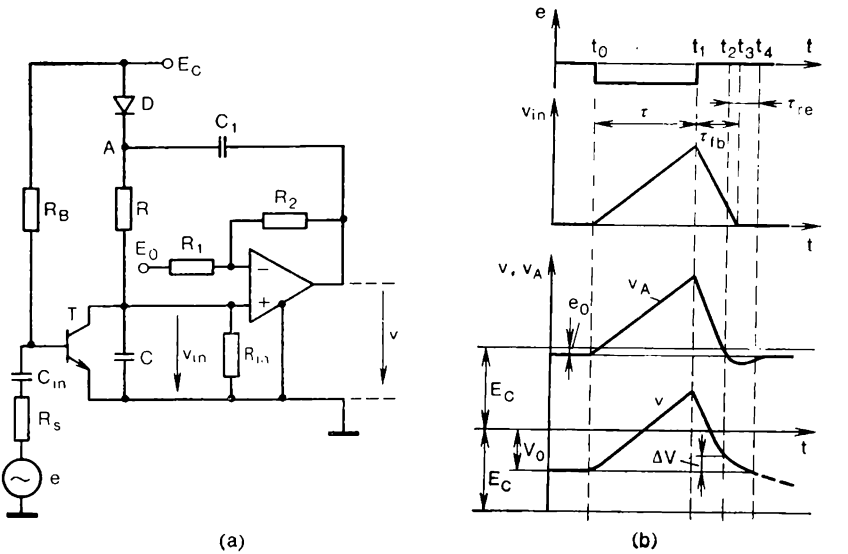


Fig. 9.16

disabled, so the charging current of the capacitor C somewhat decreases. When the cathode potential of the diode rises to $E_C - e_0$, the diode turns off and the feedback path is completed.

Now the ramp is shaped in about the same manner as examined for the generalized circuit, with the opamp having a gain (referred to the noninverting input) somewhat exceeding unity.

When the switch is turned on at time t_1, the return stroke of the ramp begins. During flyback, the capacitor C discharges via the switching transistor T in a linearized exponential manner. This is accompanied by a decrease in the output voltage and in the cathode potential of the diode. Because the initial charge on C_1 somewhat decreases during the interval from t_0 to t_2 owing to the flow of the charging current, the voltage v_A falls to the value $E_C - e_0$ towards the end of flyback (time t_2) before the charging capacitor restores its initial steady-state voltage.

At time t_2, the diode turns on and shunts the output stage (opamp), so that the opamp is caused to limit output current. At the same time, the voltage on C_1 begins to rise towards $V_{C1,0}$ which is reached at t_4 when the opamp is again driven to pass the signal linearly.

The opamp gain required to obtain the utmost in ramp linearity can be set with resistors R_1 and R_2. Owing to the additional voltage supply E_0, the initial level of output voltage can be varied and the maximum amplitude of the output ramp raised at will.

Ramp characteristics. Referring to the generalized circuit in Fig. 9.15 and recalling the opamp characteristics examined in Chap. 2, it may be taken that $R_{out} \approx 0$, the leakage resistance R_{in} of the noninverting input of the opamp is relatively high, and the opamp gain is

$$K_g \approx 1 + R_2/R_1 \tag{9.43}$$

In the circumstances, the utmost in linearity corresponding to $K = 0$ will be obtained, if

$$R_2/R_1 \approx C/C_1 + R/R_{in} \tag{9.44}$$

Quantitatively, the nonlinearity depends on how closely the equality (9.44) is satisfied, with allowance for spread in the ratings of the circuit components and R_{in}.

The initial rate of change of the ramp, $\dot{v}$ (0), is a function of the charging current

$$I_0 \approx V_{C1,0}/R \approx E_C/R \tag{9.45}$$

Because the opamp has a low output resistance, the circuit in question is practically insensitive to variations in load. The initial level of output voltage is given by

$$V_0 \approx -ER_2/R_1 \tag{9.46}$$

Flyback and recovery times. The flyback time, $\tau_{fb} = t_3 - t_1$, can be found from Eq. (9.28b), where $K_{sat} = I_B/I_{B(sat)} = E_C\beta/I_0R_B$. The recovery time τ_{re} can be found approximately, if the initial portion of the exponential voltage V is considered to be linear over the interval from t_3 to t_4 (see Fig. 9.16*b*):

$$\tau_{re} \approx \tau'_{eq}\Delta v/[\Delta v + (E_C - V_0)] \tag{9.47}$$

where $\tau'_{eq} = C_1R_L$, where R_L is the load resistance of the opamp final stage in the blocked state, and

$$\Delta v \approx \tau E_C\,[1 + 1/(K_{sat} - 1)]/C_1R$$

Parallel Bootstrap Ramp Generators

This configuration (Fig. 9.17*a*), answering the equivalent circuit of Fig. 9.14*b*, likewise has resistors R_1 and R_2 connected to the inverting input of the opamp so as to stabilize the opamp gain and to supply a current i_2 sufficient to pull up i_1 as it decreases with the decrease in the capacitor charge. If the values of R_1 through R_4 are correctly chosen, the ramp will be highly linear at both the input and output of the opamp (Fig. 9.17*b*). The initial rate of rise of v_{in} across the capacitor is controlled by I_{CO}:

$$I_{CO} = E_C/R_4 - E_0R_2/R_1R_3 \tag{9.48}$$

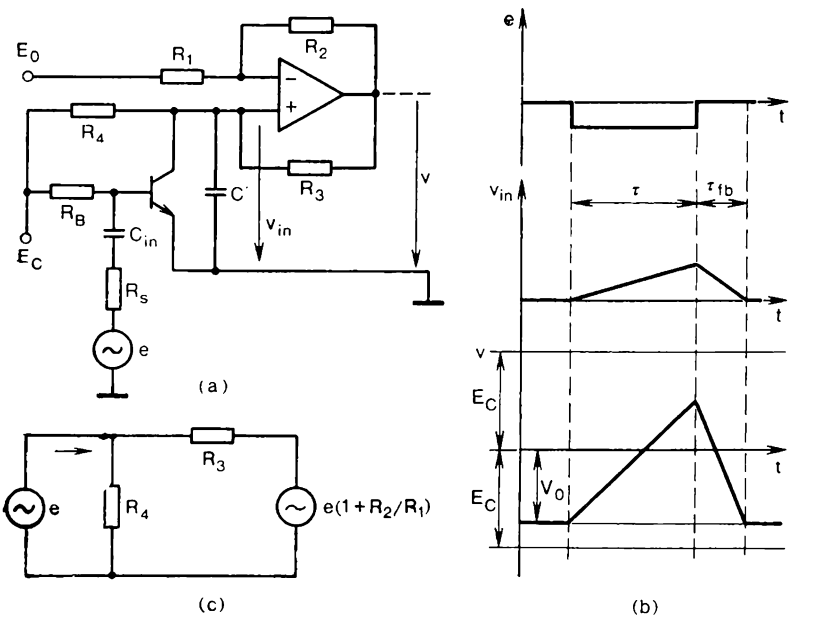


Fig. 9.17

As is seen, in contrast to the circuit of Fig. 9.16*a*, adjustment of the initial voltage level, $V_0 \approx -E_0R_2/R_1$, involves adjustment of its rate of change, which is an obvious disadvantage.

Consider the choice of values for R_1 through R_4. To begin with we write the equivalent charging time constant of the capacitor C as a product

$$\tau_{eq} = CR_{eq} \tag{9.49}$$

where R_{eq} is the equivalent incremental (differential) resistance seen looking into the noninverting input of the opamp in Fig. 9.17 during the ramp. The equivalent conductance, $G_{eq} = 1/R_{eq}$, can be found from the equivalent circuit in Fig. 9.17*c* (where the input resistances of the opamp are neglected) as

$$G_{eq} = i/e = G_4 - R_2G_3/R_1 \tag{9.50}$$

For the ramp to be perfectly linear, it is required that $G_{eq} = 0$ which corresponds to satisfying the condition

$$R_3/R_4 = R_2/R_1 \tag{9.51}$$

If we allow for the effect of the input resistances of the opamp, all of the nonlinearity will be balanced out when

$$R_2 = R_3 = R_{FB} \text{ and } R_1 = R_4 = R \tag{9.52}$$

As is shown in [27], in satisfying the condition defined by Eq. (9.52), the ratio R_{FB}/R must be chosen such that the ramp voltage has a maximum amplitude at the output of the opamp, whereas the values of the resistors depend on the expected spread in value between the input resistances of the opamp.

An advantage of the circuit in Fig. 9.17*a* over that in Fig. 9.16*a* is that it has no large-value capacitor C_1, so it is free from the associated recovery process.

Miller Integrator Ramp Generators

Effective negative feedback and freedom from voltage sources with floating terminals make the Miller integrator an excellent ramp generator especially in IC form.

The equivalent circuit of the Miller integrator ramp generator is shown in Fig. 9.18*a*. Here, the conventional transistor amplifier is represented by the equivalent circuit of a current amplifier. The supply source E_{bl} is chosen such that when the (diode or transistor) switch *Sw* is closed, the output stage of the amplifier is blocked, and the output voltage is $v = V_0 \leqslant E_s$, where V_0 is the limiting voltage level with the transistor driven at cutoff, and E_s is the supply voltage of the amplifier. When the switch is opened at time $t = t_0$, the capacitor C begins to re-charge with a current i_C which is stabilized because the amplifier input voltage which is the difference between the capacitor voltage and the amplifier output voltage, varies at a rate markedly lower than do these two voltages.

Assuming, as usual, the origin of time as $t = t_0$, the exponential change in the output voltage during the ramp may be written

$$v = v(\infty) - [v(\infty) - V_0] \exp(-t/\tau_{eq}) \qquad (9.53a)$$

where $v(\infty) = K_i R_{out} i_{in}(\infty) = K_i R_{out} E/(R + R_{in})$ and $\tau_{eq} = CR_{eq}$. The equivalent resistance, R_{eq}, in the recharging circuit

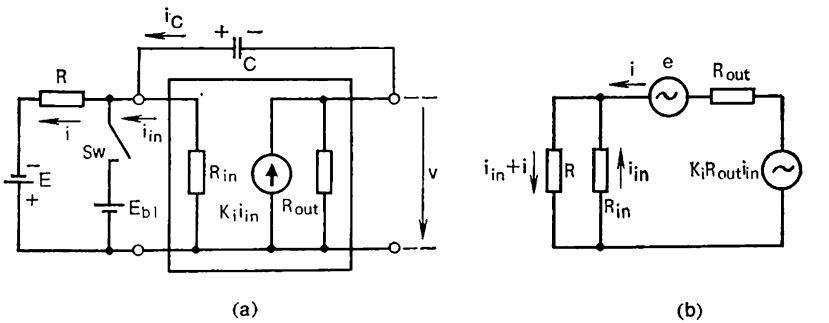


Fig. 9.18

of the capacitor can be found by placing a source of voltage e instead of the capacitor C between terminals *1* and *2* (Fig. 9.18*b*):

$$R_{eq} = e/i = \frac{(1+K_i)\,RR_{out}}{R+R_{in}} \times \left[1+\frac{R_{in}}{(1+K_i)\,R}+\frac{R_{in}}{(1+K_i)\,R_{out}}\right] \quad (9.53b)$$

Using Eqs. (9.53), we can readily determine the initial rate of change and nonlinearity of the ramp:

$$\dot{v}(0) = \frac{1}{C}\left[\frac{v(\infty)-V_0}{R_{eq}}\right] = \left[1+\frac{R_{in}}{(1+K_i)\,R}+\frac{R_{in}}{(1+K_i)\,R_{out}}\right]^{-1} \times \left[\frac{K_i}{1+K_i}\frac{E}{R}-\frac{V_0}{(1+K_i)\,R_{out}}(1+R_{in}/R)\right]\frac{1}{C} \quad (9.54a)$$

$$\xi = \tau/\tau_{eq} = \frac{\tau\,(1+R_{in}/R)}{C\,(1+K_i)\,R_{out}}\left[1+\frac{R_{in}}{(1+K_i)\,R}+\frac{R_{in}}{(1+K_i)\,R_{out}}\right]^{-1} \quad (9.54b)$$

The circuit of Fig. 9.18*a* will operate effectively, that is, it will have low nonlinearity and the initial rate of the ramp only slightly dependent on the amplifier parameters, if the following conditions are satisfied:

$$K_i \gg 1,\ \frac{R_{in}}{(1+K_i)\,R} \ll 1,\ \text{and}\ \frac{R_{in}}{(1+K_i)\,R_{out}} \ll 1 \quad (9.55)$$

In view of Eqs. (9.55), we approximately obtain

$$\dot{v}(0) \approx I_0/C,\quad \xi \approx \lambda K_E R(1+R_{in}/R)/(K_i R_{out}) \quad (9.56)$$

where $I_0 = E/R$, $\lambda = E_s/E$ and $K_E = V_m/E_s$.

The opamp examined in Chap. 2 can serve as an effective discharge-current stabilizer, if the required maximum amplitude does not exceed the linear portion of the output characteristic from $\overline{-V^L}$ to V^H.

The basic practical circuit of the Miller integrator ramp generator is shown in Fig. 9.19*a*. During the ramp time, the output resistance of the opamp is assumed to be negligible ($R_{out} \approx 0$) and the input resistance, finite ($R_{in} \neq 0$) (Fig. 9.19*b*). When the switch is open, the output voltage can be written

$$v(\infty) = -K_0 v_{in}(\infty) = -K_0 E R_{in}/(R+R_{in}) \quad (9.57a)$$

and will change exponentially with a time constant

$$\tau_{eq} = CR_{eq} \quad (9.57b)$$

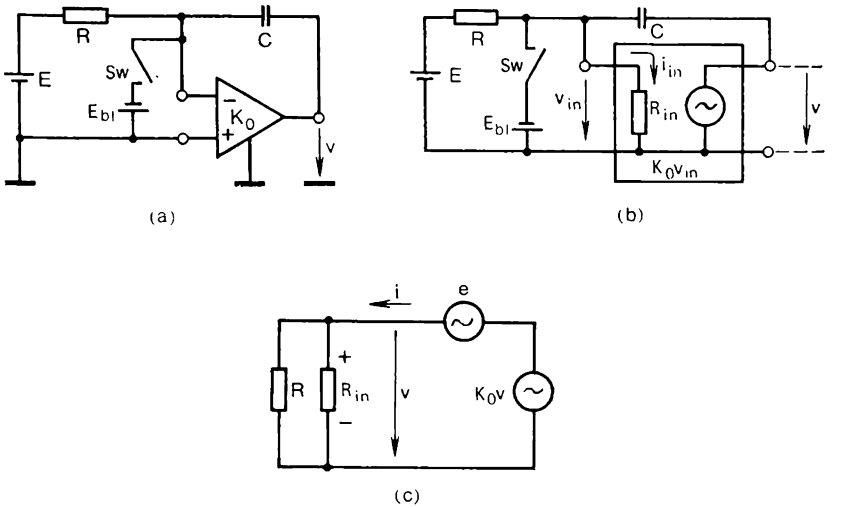


Fig. 9.19

To determine R_{eq}. we may use the equivalent circuit of Fig. 9.19*c* from which it is seen that

$$R_{eq} = v/i = (1 + K_0)\,(R \parallel R_{in}) \tag{9.57c}$$

Using Eq. (9.57a), we can determine the initial rate of change and the nonlinearity of the ramp

$$\begin{aligned} \dot{v}(0) &= -\frac{1}{C}\left[\frac{K_0}{1+K_0}\frac{E}{R} - \frac{V_0}{1+K_0}(1/R + 1/R_{in})\right] \\ \xi &= \tau/CR_{eq} = \frac{\tau}{(1+K_0)C}(1/R + 1/R_{in}) \end{aligned} \tag{9.58}$$

It is seen from Eqs. (9.58) that the circuit of Fig. 9.19*a* will operate adequately when

$$K_0 \gg 1 \text{ and } R_{in} \gg R \tag{9.59}$$

As will be recalled, the first condition is fully met within the linear portion of the amplifier characteristics, and the second limits the upper maximum value for the discharge resistor R.

When the conditions defined in Eq. (9.59) are satisfied,

$$\dot{v}(0) \approx -I_0/C \text{ and } \xi = \tau/CRK_0 = K_E\lambda/K_0 \tag{9.60}$$

Because for opamps K_0 is anywhere between 10^4 and 10^5, the Miller integrator ramp generators show a very high linearity ($\xi < 10^{-2}\%$).

Phantastrons. A *phantastron* is a relaxation oscillator in which a ramp voltage is generated across a time-controlling capacitor by

a bootstrap technique. Thus, the heart of any phantastron is any one of the negative-feedback bootstrap ramp generators examined in Sec. 9.4.

Phantastrons can be used as triggered (one-shot) and free-running relaxation oscillators and produce waveforms which compare favourably in terms of stability with those generated by other forms of multivibrator.

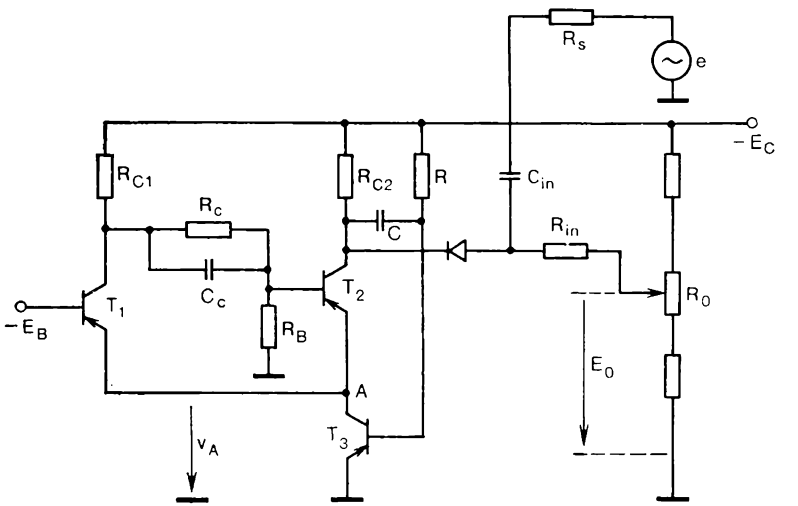
In a phantastron, the duration, period or frequency of the output pulses can be varied simply and conveniently in a linear fashion by controlling its drive voltage. Because of this, phantastrons are most frequently used as timing and adjustable pulse delay circuits. In this function, a phantastron operates as a triggered multivibrator in which the duration τ of the temporary stable state (pulse duration) is proportional to the drive voltage E_0, and its accuracy is characterized by the deviation of the actual pulse duration from the design one, $\delta\tau = \Delta\tau/\tau$.

In operation as a ramp generator, its sweep time is determined by the parameters of the generator and is independent of the trigger pulse. In this case, the important requirement is the constancy of the rate of change. The accuracy of the generator is assessed in terms of the nonlinearity and stability of the working stroke as given by the relations deduced earlier.

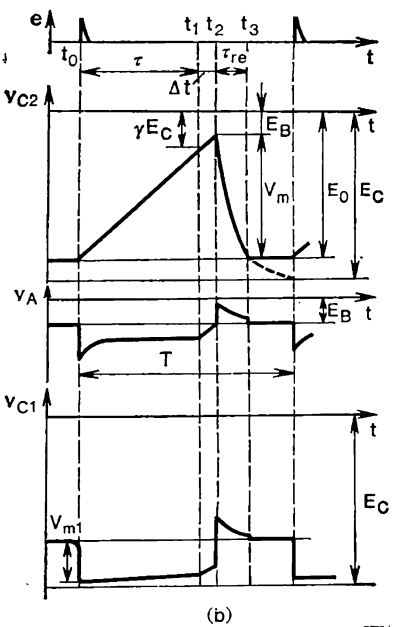
Another important characteristic of phantastrons is the maximum pulse duty factor, $\eta = \tau/T$, where T is the minimum permissible repetition period of trigger pulses (in triggered or one-shot operation) or the minimum possible period of oscillations (in free-running operation).

One of the basic circuits of a triggered transistor phantastron is shown in Fig. 9.20*a*. Initially, when the circuit is in its stable state, transistor $T2$ is OFF, whereas transistors $T1$ and $T3$ are ON. $T3$ is in the active region, and $T1$ may be in the active region or at saturation. A positive trigger current pulse is routed via a trigger diode and a capacitor to the base of $T3$ thereby bringing down its current. The potential at point A (the emitter of $T1$) is brought down, and so is the current through $T1$. The decrease in the collector potential of $T1$ is transferred via a speed-up capacitor C_c to the base of $T2$ which is thus rendered conducting and completes the positive feedback path, thereby triggering the phantastron.

The positive feedback via $T1$ is applied to two resonant circuits. In one of them, the change in the potential at point A is boosted by the collector circuit of $T1$, transferred via a coupling network, R_cC_c, to the base of $T2$, reproduced in its emitter circuit, again boosted in the collector circuit of $T1$ and so on. The other resonant circuit includes the collector circuit of $T2$. A change in the potential at point A, after it has been boosted by the collector circuit of $T1$ and transferred to the base of $T2$, is amplified by the collector cir-


R_s
e
$-E_C$
R_{C1}
R_{C2}
R
C_{in}
R_c
C
R_{in}
C_c
T_2
R_0
T_1
$-E_B$
R_B
E_0
A
v_A
T_3

(a)


e
t_1 t_2 t_3
t_0
τ
τ_{re}
t
Δt
v_{C2}
γE_C
E_B
t
V_m
E_0
E_C
v_A
E_B
t
T
v_{C1}
t
E_C
V_{m1}

(b)

Fig. 9.20

cuit of that transistor, transferred via capacitor C to the base of $T3$ to be again amplified by its collector circuit.

Negative feedback is applied to one resonant circuit containing transistors $T2$ and $T3$. The voltage across this resonant circuit stabilizes the change of potentials. For example, if the potentials at the collector of $T2$ and the base of $T3$ rise, the potential at point A will fall owing to a decrease in the current through $T3$, and the change in the emitter-to-base voltage of $T2$ will bring about a decrease in its collector current, thereby preventing any further rise in its collector potential. In this way, the resonant circuit stabilizes the current during the temporary state, whereas at the instant of switching it retards the cumulative change in the potentials, which is terminated when $T1$ is turned off by the negative change of voltage at point A.

As is seen, the rapid change (a jump, in a first approximation) in the voltages and currents is followed by a relatively slow change—this is a temporary stable state of the circuit (it extends from t_0 to t_2 in the timing diagram of Fig. 9.20*b*). The principal event during this interval is a linearized re-charging of the capacitor C, which is responsible for a linear rise in the potential at the collector of $T2$. This process can be described by the relations derived earlier, with allowance for the fact that the amplifier based on transistors $T2$ and $T3$ is a two-stage common-emitter, common-base circuit.

At time t_1, the potential at the collector of $T2$ becomes equal to that at the base, $T2$ goes into the saturation region, and the rate at which the potential at point A rises (so far very low) becomes equal to the rate of change of the ramp voltage in the collector circuit of $T2$ (which, as already noted, is fairly stable). In time Δt, when $v_A(t_2) \approx -E_B$, transistor $T1$ turns ON, the potential at the base of $T2$ rises, $T2$ moves into the active region again, and the circuit changes state, completing the flipover when $T2$ is turned OFF.

During the interval from t_2 to t_3, the charge on the capacitor C acts via the collector load resistor R_{C2} and the input resistance of transistor $T3$ to restore the circuit to its original state.

Operability conditions. The phantastron in Fig. 9.20*a* will operate adequately if its circuit parameters are chosen such that the circuit can assume two stable states (one permanent and the other temporary) and the condition for a cumulative change of state is satisfied. Initially, when the circuit is in its stable state, transistor $T2$ must be at cutoff:

$$V_{BE2} = v_{B2} - v_{E2} > 0 \tag{9.61}$$

where $v_{B2} \approx \gamma v_{C1}$, $\gamma = R_B/(R_c + R_B)$, and $v_{E2} \approx -E_B$ (neglecting the voltage drop across r_{in1}).

The value of v_{C1} depends on the region in which the ON transistor $T1$ is operating. If it is at saturation,

$$\alpha_1 I_{E1} \geqslant I_{C1(sat)} \tag{9.62a}$$

where

$$I_{C1(sat)} = (E_C - E_B)/R_{C1}$$

$$I_{E1} \equiv I_{C3} \approx \beta_3 E_C/R$$

Solving Eq. (9.62a) for R, we obtain the condition for $T1$ to be at saturation in the following form

$$R \leqslant R_{sat} = \alpha_1 \beta_3 R_{C1}/(1 - E_B/E_C) \tag{9.62b}$$

If $R \leqslant R_{sat}$, v_{C1} will be approximately equal to $-E_B$, and transistor $T2$ will be turned OFF, provided

$$\gamma < 1 \tag{9.63a}$$

If $R > R_{sat}$, transistor $T1$ will be in the active region, and its collector potential will be

$$v_{C1} \approx -E_C + I_{C1}R_{C1} = -E_C + \alpha_1 \beta_3 R_{C1} E_{C1}/R$$

where no allowance is made for the shunting effect of the divider $R_c + R_B$ which exceeds R_{C1} in value. Substituting the above expression for v_{C1} in Eq. (9.61) gives the condition for $T2$ to turn off:

$$\gamma \leqslant \gamma_{max} = E_B/E_C\,(1 - \alpha_1 \beta_3 R_{C1}/R) \tag{9.63b}$$

In the temporary state, it is essential to maintain the ramp voltage in the collector circuit of $T2$ highly linear and to have it control the duration τ of the temporary state.

In accord with Eqs. (9.55) defining the conditions for the correct operation of a negative-feedback ramp generator, the parameters of the circuit in Fig. 9.20*a* must be as follows:

$$\begin{gathered} \beta_3 \alpha_2 \gg 1 \\ r_{in3}/\beta_3 R \ll 1 \\ r_{in3}/\beta_3 R_{C2} \parallel r_{out2} \ll 1 \end{gathered} \tag{9.64}$$

The voltage $v_{C2}(t)$ will be time-controlling if transistor $T2$ moves into the saturation region before transistor $T1$ turns ON, that is

$$v_{BE1} = v_{B1} - v_{E1} > 0 \tag{9.65a}$$

where

$$v_{B1} = -E_B$$

$$v_{E1}(t_1) \approx v_{B2}(t_1) \approx -\gamma E_C + i_{B2}(t_1)\,(R_c \parallel R_B)$$

If the current gain of $T2$ is sufficiently high (which usually is the case), the second term in the expression for v_{E1} is negligibly small, and we may write

$$v_{E1}(t_1) = V_{E1} \approx -\gamma E_C$$

Then Eq. (9.65a) yields the condition for $T2$ to become saturated before $T1$ turns ON in the following form:

$$\gamma > \gamma_{min} = E_B/E_C \tag{9.65b}$$

The final condition is that the positive feedback circuit be sufficiently effective, even though all the three transistors should move into the active region simultaneously. This is essential for the phantastron to change state at a cumulatively increasing rate at the start and end of the temporary state. This condition is satisfied with $K_{FB} > 1$, where K_{FB} is the loop gain on opening the path common to both resonant circuits to which positive feedback is applied (this may be the connection from the collector of $T1$ to the base of $T2$). Analysis shows that this condition is satisfied, if the equivalent load in the collector of $T1$ (with C_c deemed shorted out) is not too small:

$$R_{C1} \parallel R_B > r_{in1}/(1+\beta) \tag{9.66}$$

Pulse duration. The duration of the temporary state or, which is the same, the duration of the pulse generated by a phantastron ($\tau = t_0 - t_2$) can readily be calculated on assuming the ramp voltage perfectly linear:

$$\tau \approx \frac{E_0 - E_B}{\dot{v}(0)} = \tau_0 (E_0' - R_B') \tag{9.67}$$

where $\tau_0 = RC$, $E_0' = E_0/E_C$, and $E_B' = E_B/E_C$.

It follows from Eq. (9.67) that to a first approximation the pulse duration is directly proportional to the control voltage E_0. Therefore, the scale of the linear potentiometer r_0 may be calibrated directly in units of time—micro- or milliseconds. Adjustment is possible within the interval equal to the pulse rise time, $\tau_r = t_1 - t_0$, and the degree of linearity can be evaluated in the following manner. Using the generalized exponential equation for $v_{C2} = v$, we obtain

$$\tau_r = -\tau_{eq} \ln\left[1 - \frac{v(\tau_r) - v(0)}{v(\infty) - v(0)}\right] \tag{9.68a}$$

where

$$\tau_{eq} \approx CR_{C1}(1 + \beta_3\alpha_2)/(1 + r_{in1}/R)$$

For $R_{C1} \ll r_{in2}$,

$$v(\tau_r) \approx -\gamma E_C$$

and

$$v(0) \approx -E_0$$

On expanding Eq. (9.68a) into a descending power series and retaining only the quadratic term, we get

$$\begin{aligned}\tau_r &\approx \tau_{eq}\frac{v(\tau_r)-v(0)}{v(\infty)-v(0)}\left[1+\frac{1}{2}\frac{v(\tau_r)-v(0)}{v(\infty)-v(0)}\right]\\ &=\frac{E_0-\gamma E_C}{\dot{v}(0)}\left[1+\frac{1}{2}\frac{E_0-\gamma E_C}{\tau_{eq}\dot{v}(0)}\right]\\ &=\tau_0(E_0'-\gamma)\left[1+\frac{1}{2}\frac{E_0'-\gamma}{\beta_3}\frac{R}{R_{C2}}\right]\end{aligned} \tag{9.68b}$$

The shape of the parabolic function $\tau_r = f(E_0')$ is illustrated in Fig. 9.21. The maximum absolute deviation $\Delta\tau_r$ and the maximum fractional deviation $\delta\tau_r = \Delta\tau_r/(E'_{0,max} - E'_{0,min})$, of the pulse duration from the linear calibration of the control potentiometer (r_0 in Fig. 9.20*a*) can be found, using the same technique as was used to find the offset error of the ramp voltage in Sec. 9.2. More specifically,

$$\delta\tau_r = \frac{1}{8}\frac{E'_{0,max}-E'_{0,min}}{\tau_{eq}\dot{v}(0)} = (1/8\beta_3)(R/R_{C2})(E'_{0,max}-E'_{0,min}) \tag{9.69}$$

It is seen from Eq. (9.69) that the linearity of control improves with decreasing range $E'_{0,max} - E'_{0,min}$, and with decreasing ratio R/R_{C2}. The minimum permissible value of R/R_{C2} is limited by the maximum permissible flyback time, $T - \tau_{r,max}$ (assuming that the decay time $\Delta t = t_2 - t_1$ is negligible in comparison with $\tau_{r,max}$, which usually is the case in practice).

Assuming the maximum pulse duration, $\tau_{r,max}$, the flyback time τ_{fb} is given by

$$\tau_{fb} = k_{re}CR_{C2} = k_{re}R_{C2}\tau_0/R \tag{9.70}$$

where $k_{re} = \ln[(1 - E'_B)/(1 - E'_{0,max})]$.

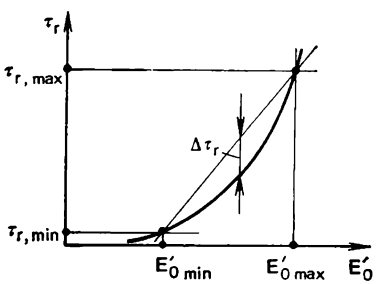


Fig. 9.21

From Eq. (9.70) it follows that

$$(R/R_{C2})_{min} = \frac{k_{re}\tau_0}{T-\tau_{r,max}} = \frac{k_{re}}{T-\tau_{r,max}} \frac{\tau_{r,max}}{E'_0-\gamma} = \frac{\eta}{1-\eta} \frac{k_{re}}{E'_{0,max}-\gamma} \quad (9.71)$$

On substituting for $(R/R_{C2})_{min}$ from Eq. (9.71) in Eq. (9.69), we get

$$\delta\tau_{r,min} = \frac{1}{8} \frac{k_{re}\eta}{\beta_3(1-\eta)} \frac{E'_{0,max}-E'_{0,min}}{E'_{0,max}-\gamma} \quad (9.72a)$$

If $E'_{0,min}$ is chosen to be approximately equal to γ (which usually is the case in practice), then

$$\delta\tau_{r,min} \approx \frac{1}{8} \frac{k_{re}\eta}{\beta_3(1-\eta)} \quad (9.72b)$$

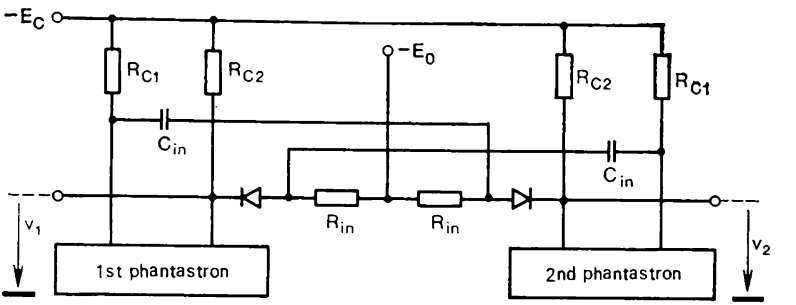
The function defined by Eq. (9.72b) is analogous to the function $\xi_{min} = f(\eta)$ in the types of ramp generator examined earlier. Adjustment of pulse duration can be made linear, using the same techniques as are applied to Miller integrator ramp generators (to switch load in or out, or to use high-gain amplifiers).

Output pulse amplitude. When a phantastron is used as a source of roughly rectangular pulses, its load is usually placed in the collector circuit of $T1$. In such cases, the stability of output pulse amplitude can be improved by operating $T1$ at saturation in the initial stable state. Then,

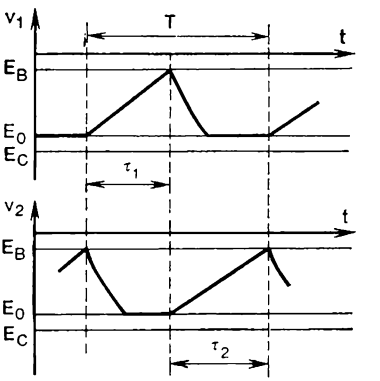
$$V_m = |V_{max}| - |V_{min}| = E_C - E_B - I_{CO}R_{C1}$$

Free-running operation. If, in the circuit of Fig. 9.20a, we remove R_c and return the lower terminal of R_B to the negative side of the power supply, the initial stable state will be stable only temporarily. Its duration is determined by the exponential discharge of the capacitor C_c via the base resistor R_B and the collector circuit of transistor $T1$. After $T2$ turns ON, the output pulse is shaped in exactly the same manner as in a triggered phantastron, by the linear discharge of the capacitor C.

As regards the stability of the period of free oscillations, such a phantastron has no advantage over multivibrators. A better frequency stability, similar to that of a triggered phantastron, is offered by the circuit of Fig. 9.22*a*. It is a combination of two cross-coupled triggered phantastrons which are triggered in turn at the end of each other's temporary state (Fig. 9.22*b*). If E_0 is set to the same value for either phantastron, the circuit will maintain the period of oscillations as a linear function of the drive voltage to a high degree of accuracy. Linear adjustment for the frequency of oscillation can be obtained by applying a controllable voltage to the time-setting resistors R in the base circuits of transistors $T3$.



(a)



(b)

Fig. 9.22

REVIEW QUESTIONS AND PROBLEMS

9.1. List the main characteristics of a ramp generator and derive the equations defining the initial rate of change, slope error, and offset error of the ramp voltage.

9.2. Draw the circuit of a simple CR ramp generator; determine the nonlinearity and stability of the ramp voltage, and evaluate the effect of load.

9.3. Draw the circuit and explain the operation of a ramp generator using a diode switch; determine the flyback time.

9.4. Draw the circuit and explain the operation of a ramp generator using a transistor switch; determine the turn-off delay time of the switch and the flyback time of the ramp (for a saturated and nonsaturated switch).

9.5. Draw the circuit of a transistor constant-current ramp generator and show the waveform of its output signal; explain its principle of operation.

9.6. Draw an equivalent circuit for the ramp generator of Fig. 9.11; determine the initial rate and linearity of the ramp and the effect of load.

9.7. Optimize the circuit parameters of the ramp generator shown in Fig. 9.11 for maximum linearity.

9.8. Using appropriate equivalent circuits, explain how the nonlinearity of the ramp voltage can be minimized by bootstrapping.

9.9. Draw a generalized circuit for a series bootstrap ramp generator; formulate the requirements for the choice of its circuit parameters; determine the initial rate and nonlinearity of the ramp voltage.

9.10. Draw the circuit of a series bootstrap ramp generator using an IC opamp, plot its waveform, explain the principle of operation, determine the ramp characteristics, flyback time, and recovery time.

9.11. Draw the circuit of a parallel bootstrap ramp generator, plot its waveform, and select its circuit parameters for maximum linearity.

9.12. Draw the basic circuit of the Miller integrator ramp generator; explain its operation, determine the initial rate of change and nonlinearity of the ramp, with the circuit parameters chosen to meet the conditions for operability.

9.13. Draw the circuit and plot the waveform of a single-stage Miller integrator ramp generator; explain its operation.

9.14. Determine the ramp characteristics and the flyback time for the generator of Problem 9.12.

9.15. Derive equations defining the effect of the resistive and capacitive load components on the ramp voltage of Problem 9.12.

9.16. Select the circuit parameters for a ramp generator with a nonswitched and switched load for maximum linearity.

9.17. Draw the circuit of an opamp ramp generator, explain its operation, determine the initial rate of change and nonlinearity of the ramp voltage, with the circuit parameters chosen to satisfy the conditions for operability.

9.18. Draw the circuit of a phantastron, plot the waveform of its output signal, and explain its operation.

9.19. Define the conditions for the correct operation of the phantastron in Fig. 9.20.

9.20. Determine the pulse duration and linearity of duration adjustment for the phantastron of Fig. 9.20.

9.21. Draw the circuit of a free-running phantastron, show the waveform of its output signal, and explain its operation.

9.22. Draw the circuit of a pulse generator made up of two cross-coupled phantastrons, plot the waveform of output pulses, and explain its operation.

10 SYNCHRONIZATION OF RELAXATION OSCILLATORS. PULSE FREQUENCY DIVISION

10.1. GENERAL

A free-running relaxation oscillator is said to be operating synchronized when it is locked in step with an external period signal, usually called a *synchronizing signal.* Whereas in free-running operation, the frequency of a relaxation oscillator is decided by its circuit parameters, in synchronized operation (subject to certain conditions) it is equal to or is a multiple of the frequency of the synchronizing source. The ratio of the sync frequency to the output frequency of a synchronized relaxation oscillator may be called the *division* or *synchronization ratio*, n_{sync}. It may range from unity to several thousand. When $n_{sync} = 1$, the relaxation oscillator is operating at its free-running or natural frequency. When n_{sync} exceeds unity but remains an integer, the relaxation oscillator operates at a submultiple of the free-running frequency, thereby functioning as a pulse-repetition rate or pulse-repetition frequency (PRR or PRF) divider.

A triggered relaxation oscillator normally operates in synchronism with its trigger-pulse source, but, given certain conditions, it can also operate as a pulse repetition rate divider. However, output pulses will cease to appear as soon as the trigger source stops supplying trigger pulses. In contrast, removal of trigger pulses will leave a free-running relaxation oscillator operating at its own frequency.

As a rule, the sync voltage has the form of short pulses (in comparison with pulse repetition period or spacing) or a sine wave. In practice, the most commonly used sync pulses have a fast rise time, because they ensure the best time lock between the relaxation oscillator and the sync source.

The output pulses of a frequency divider depend on its circuit configuration and can be shortened, if necessary.

PRR (or PRF) division is widely utilized to lock relaxation oscillators to one another in a multichannel system, to measure or maintain time intervals in radar and radio navigation systems, to supply precise time marks in oscilloscopes for observation of periodic processes, to generate sweep waveforms in TV, etc.

10.2. SYNCHRONIZATION OF FREE-RUNNING RELAXATION OSCILLATORS

Synchronization bands. Consider the basic relations that apply to synchronized relaxation oscillators, using as an example a blocking oscillator synchronized by short pulses impressed on an additional transformer winding (Fig. 10.1*a*). The manner in which this is done is illustrated in Fig. 10.1*b*. The figure shows the voltage waveforms across the additional winding and at the transistor base at the end of two cycles of free-running operation. As is seen, the period T_{sync} is many times the pulse duration τ, which is always the case in practice. Referring to Fig. 10.1*b*, when the time-controlling voltage v_C is combined with the sync pulses applied to the base winding of the transformer, the resultant emitter-to-base voltage v_{BE} is brought down, and the transistor turns on when this voltage becomes equal to the threshold value, V_{th}, which occurs at every second sync pulse ($n_{sync} = T/T_{sync} = 2$). Thus, a necessary, but not sufficient condition for operation at the synchronization ratio n_{sync} is

$$T_0 \geqslant T = n_{sync} T_{sync} \tag{10.1}$$

where T_0 is the free-running or natural period of the relaxation oscillator.

In selecting the circuit parameters required to obtain the desired synchronization ratio, let us deem, at least approximately, that the timing voltage is linear and that the duration of the sync pulse and of blocking-oscillator pulses is negligible. The waveform idealized in this sense, within one cycle for $n_{sync} = 4$, is shown in Fig. 10.2*a*.

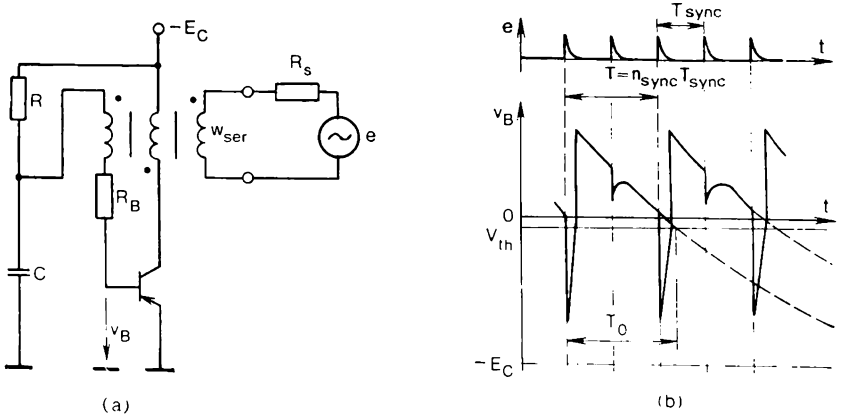


Fig. 10.1

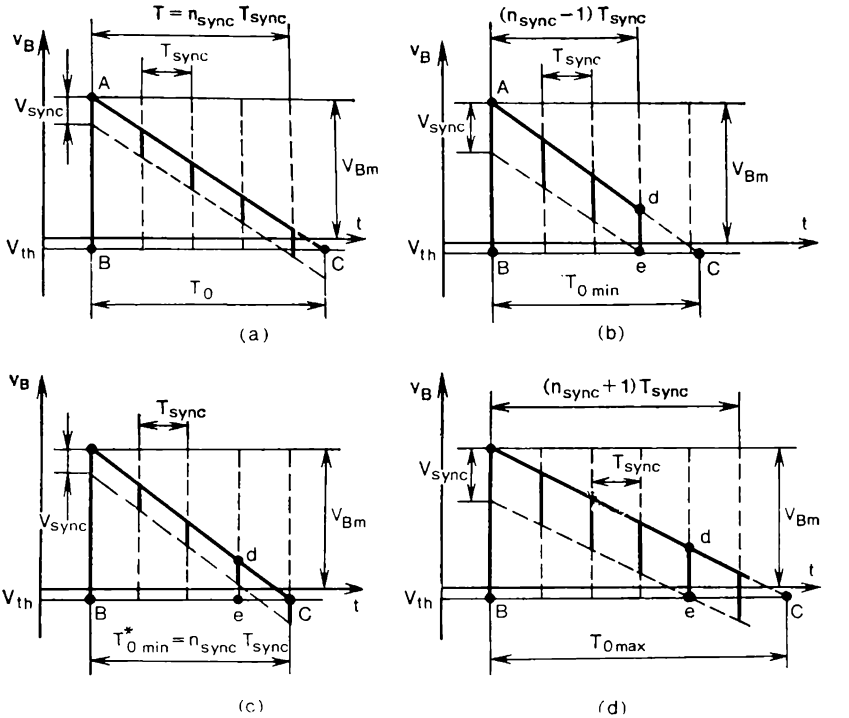


Fig. 10.2

Proper synchronization can be upset if the free-running period becomes shorter than $n_{sync}T_{sync}$. This can happen in any one of two cases. Firstly, if the sync pulses are sufficiently large in amplitude, a decrease in the free-running period or an increase in the amplitude of the sync pulses will cause the transistor to turn ON after the (n — 1)st sync pulses. Geometrically, the free-running period has a minimum value, $T_{0,min}$, when the $(n - 1)$st sync pulse occurs at the point e (Fig. 10.2*b*). From the similarity of the triangles ABC and deC we may write

$$\lambda = 1 - (n_{sync} - 1)\,\theta_{max} \tag{10.2}$$

where $\lambda = V_{sync}/(V_{B,m} - V_{th})$ is the relative amplitude of sync pulses and $\theta = T/T_0$ is the relative period of oscillation.

Secondly, if the sync pulses are not too large in amplitude and the natural period is $T_0 \leqslant n_{sync}T_{sync}$ the oscillations will fail to lock in step with the sync pulses and will cease within the nth cycle of the sync pulses. Geometrically, the minimum value, $T^*_{0,min}$

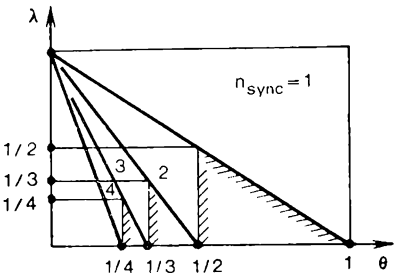


Fig. 10.3

occurs where the timing-voltage line cuts the threshold voltage line (Fig. 10.2*c*):

$$T^*_{0,\min} = n_{sync} T_{sync} \quad \text{or} \quad \theta^*_{max} = 1/n_{sync} \tag{10.3}$$

Synchronized operation can also be upset if the natural period exceeds $n_{sync}T_{sync}$ or because of a decrease in the amplitude of sync pulses. The value of $T_{0,max}$ can be found geometrically at the point where the nth sync pulse cuts the threshold voltage line (Fig. 10.2*d*):

$$\lambda = 1 - n_{sync}\theta_{min} \tag{10.4}$$

where $\theta_{min} = T_{sync}/T_{0,max}$.

Using Eqs. (10.2) through (10.4), we can readily plot a diagram of synchronization bands for various synchronization ratios. These equations are in effect the equations of straight lines delineating bands with specified values of n_{sync}. Referring to the plot of Fig. 10.3, we can determine the nominal value for the natural period of a blocking oscillator, the amplitude of sync pulses for a given value of n_{sync}, and the maximum deviations in these quantities which leave the desired form of synchronization. Referring to the diagram, an increase in the synchronization ratio causes the respective region of stabilization to become narrower. Because of this, more stringent requirements have to be met as regards the stability of θ and λ.

When the trigger pulses have a finite duration, the regions of constant n_{sync} are separated in the diagram by spaces corresponding to intersections between the V_{th} line and the peaks of the trigger pulses [1]. Analysis shows that the optimal amplitude of sync pulses is approximately equal to the increment ΔV in the linear time-controlling voltage over the repetition period T_{sync} of the sync pulses. For transistor blocking oscillators, n_{sync} usually ranges from 3 to 5 and is limited by the temperature instability of T_0.

The synchronization mechanism and the basic relations examined above apply as well to other free-running relaxation oscillators, notably multivibrators [4] and phantastrons (on the assumption that the synchronizing voltage is linear). If no special measures are taken to stabilize the repetition period of a multivibrator, its $n_{sync,\ max}$ will be 3 to 5, as for blocking oscillators.

10.3. PULSE RATE DIVISION BY TRIGGERED RELAXATION OSCILLATORS

In free-running relaxation oscillators, sync pulses terminate the temporary state, thereby limiting the half-period of oscillation. By contrast, in triggered relaxation oscillators (multivibrators and phantastrons) used as PRF dividers, sync pulses mark the beginning of the temporary state. Given certain conditions to be defined later, it is possible to make n_{sync} independent of the amplitude (peak value) V_{sync} of the trigger pulses or, if V_{sync} be held constant, to obtain substantially higher values of n_{sync}.

Consider the operation of a direct-coupled one-shot with a cutoff diode $D2$ (Fig. 10.4*a*), used to divide the recurrence rate of trigger pulses by a factor n_{sync}. The trigger pulses are applied to the collector of $T1$ via a trigger diode $D1$. If at $t < t_0$ (Fig. 10.4*b*) the multivibrator is in its initial stable state, the trigger pulse arriving at time t_0 passes through $D1$ and starts up the multivibrator. During the temporary state (the interval τ) and the greater part of the recovery time (the interval τ'_{re}) the diode $D1$ is reverse-biased by a voltage exceeding the peak value, V_{sync}, of the trigger pulses. The

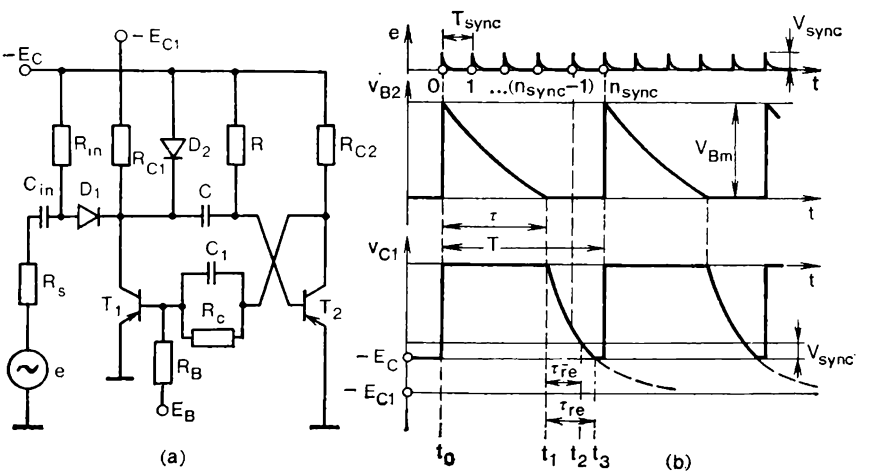


Fig. 10.4

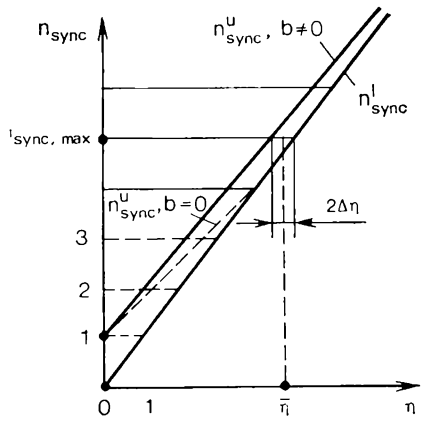


Fig. 10.5

multivibrator is rendered insensitive not only to the pulses arriving during the temporary state (these pulses cannot change the state of the circuit even if their amplitude exceeded the reverse bias on the diode, $V_{sync} > E_C$), but also to the pulses arriving during the interval from t_1 to t_2. The multivibrator is again started by the trigger pulse arriving first after time t_3, when the reverse bias on the diode falls below V_{sync}. In this way, the rate at which the multivibrator swings back and forth is $1/n_{sync}$ of the trigger pulse rate (in Fig. 10.4 *b*, $n_{sync} = 5$).

Consider the conditions for a relaxation oscillator to operate as a PRF divider when the repetition period T_{sync} of sync pulses is shorter than the recovery time, τ_{re}. The division mode will be effected, if the trigger diode is turned on by the nth sync pulse when the circuit has returned to its original state, rather than by the $(n-1)$st pulse. By inspection of the circuit in Fig. 10.4*b*, the applicable conditions may be written as:

$$\begin{aligned} (n_{sync}-1)\,T_{sync} &\leqslant \tau + \tau'_{re} \\ n_{sync}T_{sync} &\geqslant \tau + \tau_{re} \end{aligned} \tag{10.5}$$

where

$$\tau_{re} = -R_{C1}C \ln (1 - E_C/E_{C1})$$
$$\tau'_{re} = -R_{C1}C \ln [1 - (E_C - V_{sync})/E_{C1}]$$

From Eq. (10.5), we get

$$n^{H}_{sync} > n_{sync} > n^{L}_{sync} \tag{10.6}$$

where

$$n^{H}_{sync} = \eta(1 + bk_{re}) + 1$$
$$n^{L}_{sync} = \eta(1 + k_{re})$$
$$\eta = \tau/T_{sync}$$
$$b = \tau'_{re}/\tau_{re}$$
$$k_{re} = \tau_{re}/\tau$$

The straight lines representing the functions defined by Eqs. (10.6) are shown in Fig. 10.5. As is seen, an increase in n_{sync} entails a decrease in the maximum deviation, $\pm\Delta\eta$, of η from the average optimal value, $\bar{\eta}$, and, as a consequence, more stringent requirements as

regards the stability of pulse duration τ for the multivibrator and the amplitude of trigger pulses. Using Eqs. (10.6) and referring to Fig. 10.5, the relation between the maximum attainable division ratio, $n_{\text{sync, max}}$, and the circuit parameters may be written

$$n_{\text{sync, max}} = \frac{1+k_{re}}{1+k_{re}-(1-\delta\eta)(1+bk_{re})/(1+\delta\eta)} \approx \frac{1+k_{re}}{k_{re}(1-b)+2\delta\eta} \qquad (10.7)$$

where $\delta\eta = \Delta\eta/\bar{\eta}$.

If E_C/E_{C1} is less than 0.5, then $b \approx 1 - \lambda$, where $\lambda = V_{\text{sync}}/E_C = V_{\text{sync}}/V_{0,m}$, and Eq. (10.7) may be simplified

$$n_{\text{sync, max}} = (1+k_{re})/(\lambda k_{re}+2\delta\eta) \qquad (10.8)$$

It follows from Eq. (10.8) that an increase in the relative amplitude of trigger pulses and in the instability of the temporary state (the parameter $\delta\eta$) entails a decrease in $n_{\text{sync, max}}$. When λ exceeds $2\delta\eta$, there is a decrease in $n_{\text{sync, max}}$ with the recovery time (the parameter k_{re}), whereas when λ is less than $2\delta\eta$, there is an increase in $n_{\text{sync, max}}$ with the recovery time. The latter case is more typical of transistor relaxation oscillators in which the temperature instability of τ is especially pronounced. If $\delta\eta = 0.15$, $\lambda = 0.1$, and $k_{re} = 0.1$ and 1, then $n_{\text{sync, max}}$ will be 3 and 5, respectively.

To minimize the effect of the trigger-pulse amplitude, it is essential that the temporary state be terminated between the $(n-1)$st and the nth trigger pulses. Then $n_{\text{sync, max}}$ will be given by Eq. (10.8) for $\lambda = 1$.

In practice, individual PRF dividers based on the principles examined above are combined into more elaborate systems capable of higher values of $n_{\text{sync, max}}$, better adjustment capabilities, and higher reliability. Among them are frequency dividers based on pulse discrimination and feedback chain dividers [4, 38].

Another way to divide the repetition rate of pulses is by means of a digital counter (see Chap. 13) which generates an overflow pulse in response to each μth input pulse, where μ is the modulus of counting.

REVIEW QUESTIONS AND PROBLEMS

10.1. Plot synchronization diagrams for the blocking oscillator of Fig. 10.1, assuming normal synchronization by the n_sth pulse; consider the conditions in which this form of synchronization will be disturbed; plot appropriate timing diagrams.

10.2. Plot a diagram of synchronization bands for several values of n_{sync}.

10.3. Using a diagram of synchronization bands, establish the relation between the relative deviations in $\delta\lambda = \Delta\lambda/\lambda_{\text{opt}}$ and $\delta\theta$

$= \Delta\theta/\theta_{opt}$, on the one hand, and the maximum synchronization ratio, $n_{sync,\,max}$, on the other. In working out the problem, assume that λ_{opt} and θ_{opt} are the coordinates of the centre point of the dispersion triangle with sides $2\Delta\lambda$ and $2\Delta\theta$, lying in the range of values for $n_{sync,\,max}$ so that one of its corners lies on the straight line defined by Eq. (10.4) and the opposite vertical side coincides with the straight line defined by Eq. (10.3).

10.4. Using the results of Problem 10.3, show that the optimal amplitude of sync pulses is approximately equal to the change ΔV in the linear control voltage over the repetition period of sync pulses.

10.5. Draw the circuit of a one-shot multivibrator operating as PRF divider; plot the waveform of its output signal and explain its operation.

10.6. Derive Eq. (10.7), and determine $n_{sync,\,max}$ independent of the amplitude of trigger pulses.

11 GENERATION OF SAWTOOTH (RAMP) CURRENT

11.1. GENERAL

In magnetically deflected cathode-ray tubes the deflection of the electron beam is proportional to the current in the deflection coils (often collectively called the deflection yoke); hence, a sawtooth (ramp) waveform of current is required.

The major difficulties associated with the generation of sawtooth (ramp) current are to avoid or eliminate the distortion produced in the deflection waveform by transients at the start or end of the sweep and the nonlinearities and instabilities in the tube or transistor parameters. These difficulties can be resolved in any one of several ways, depending on the specific requirements for the accuracy, speed, economy and other features of the deflection system.

The sawtooth current generators used in radar, where the accuracy requirements are especially stringent may be classed into two basic groups. Those in the first are in effect a combination of a suitable sawtooth voltage generator and a current amplifier which converts the generated voltage into the deflecting current. As will be shown later, an appropriate amount of negative feedback must be applied to the current amplifier so as to minimize frequency and nonlinear distortion in the deflection current. Those in the other group are basically opamp integrators with a feedback voltage derived from across the deflecting coil placed in the output circuit of the final amplifier stage.

The sawtooth current generators used as TV timebases where the accuracy requirements are more lenient usually have no negative feedback amplifiers, and the amplitude of the deflecting current is boosted by the transients occurring at the beginning of fly-back.

For each particular type of the deflection system, there is a certain definite number of ampere-turns, $I_\tau w$, required to deflect the electron beam across the CRT screen. Here, I_τ is the maximum required deflecting current, and w is the number of turns in the deflecting coils.

11.2. TRANSIENT RESPONSE OF THE DEFLECTION SYSTEM

Consider a linear equivalent circuit of the deflection system (Fig. 11.1) where r and C are the ohmic resistance and the stray capacitance of the system, respectively, $R_{eq} = R_{out} \parallel R_{sh}$ is the equivalent shunting resistance which includes the internal resistance of the output stage and the damping resistance R_{sh}. Assume that the scan portion (working stroke) of the sweep current is linear:

$$i = A_i t$$

where $0 < t \leqslant \tau$ (where τ is the scan time of the linear sweep). Then the voltage across the deflection coils, $v(t)$, and the current of the output (or control) stage, $i_0(t)$, may be written

$$v(t) = LA_i + rA_i t \tag{11.1a}$$

$$i_0(t) = P(t) + B_0 + A_0 t \tag{11.1b}$$

where $P(t) = \delta(t)\, CLA_i$ is the initial delta function of the control current, $\delta(t)$ is the impulse function, $B_0 = L/R_{eq} + Cr$ is the current step, and $A_0 = A_i(1 + r/R_{eq})$ is the rate of change of the deflecting current.

For real deflection systems, it is usual that

$$Cr \ll L/R_{eq} \text{ and } r/R_{eq} \ll 1 \tag{11.2}$$

If so, the relation between B_0 and A_0, on the one hand, and A_i, on the other, has the form

$$B_0 \approx A_i L/R_{eq}, \quad A_0 \approx A_i \tag{11.3}$$

Thus, for the deflecting current to rise linearly with time, it is essential that the output-stage current should contain an initial impulse of an infinite amplitude so that the charge $Q_0 = LC_s A_i$ on the stray capacitance can build up instantaneously, and the current step must be referenced to the rate of linear change of the current:

$$B_0 \approx A_0 L/R_{eq} \tag{11.4}$$

Let us see how the waveform of the deflecting current will be distorted, if the function $i_0(t)$ is other than defined by Eq. (11.1b).

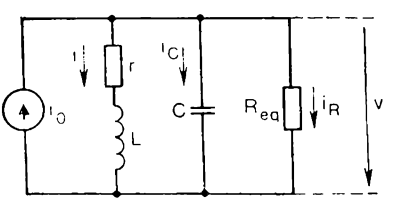


Fig. 11.1

This calls for knowledge of the initial current in the deflection system. The Laplace transforms of the deflection-system voltage and current are

$$i(s) = \frac{i_0(s)}{CL(s-s_1)(s-s_2)} + i_{\text{init}}(s) \tag{11.5a}$$

$$v(s) = i(s)(sL + r) - Li(0) \tag{11.5b}$$

where

$$i_{\text{init}}(s) = \frac{v(0)/L + i(0)/CR_C + si(0)}{(s-s_1)(s-s_2)}$$

and s_1 and s_2 are the roots of the characteristic equation

$$s^2 + s(1/CR_{eq} + r/L) + (1 + r/R_{eq})/CL = 0 \tag{11.6}$$

In view of Eqs. (11.2), we obtain from Eq. (11.6) that

$$s_{1,2} = (1 \pm \sqrt{\Delta})/2CR_{eq} \tag{11.7}$$

where $\Delta = 1 - 1/\gamma^2$, $\gamma = \rho/2R_{eq}$, and $\rho = \sqrt{L/C}$.

To begin with, let us see how the sweep current waveform is produced by a trapezoidal control current, $i_0(t) = A_0t + B_0$. Two forms of the transient response are of interest, namely: the critically damped response when negative *current* feedback is applied to the sweep-generator amplifier, and the overdamped (aperiodic) response when negative *voltage* feedback is applied. The first case, for which $\Delta = 0$ and $\gamma = 1$, takes place when the equivalent shunting resistance is

$$R_{eq} = R_{eq,\,cr} = \rho/2 \tag{11.8}$$

Under zero initial conditions, the deflecting current will obey the following law:

$$i(t) = A_0t\{1 + (1 + B_0/A_0\tau_0)\exp(-t/\tau_0) + B_0(1 - 2A_0\tau_0)/B_0)[1 - \exp(-t/\tau_0)]\} \tag{11.9}$$

where $\tau_0 \approx \sqrt{CL} = 2CR_{eq,\,cr}$.

With any value of B_0, the initial portion of the sweep waveform is nonlinear, and the change in the deflecting current will be linear only after the transients described by the exponential term in Eq. (11.9) have died out.

The current step value determines the time shift in the linear portion of the sweep waveform. With $B_0 < 2A_0\tau_0$, this shift is positive, and the linear portion of the actual sweep waveform lags behind the ideal sweep current, $i = A_it$. With $B_0 > 2A_0\tau_0$, the real sweep leads the ideal one.

The time shift is fully balanced out when

$$B_0 = 2A_0\tau_0 \tag{11.10}$$

Then the voltage across and the current in the deflection system (Fig. 11.2) are

$$v(t) = LA_i [1 - \exp(-t/\tau_0)] + LA_i (t/\tau_0) \exp(-t/\tau_0) \quad (11.11a)$$

$$i(t) = A_i t [1 - \exp(-t/\tau_0)] \quad (11.11b)$$

where $A_i = A_0$.

The maximum per cent slope error, δi, in the function defined by Eq. (11.11b) which occurs at $t = \tau$, is given by

$$\text{per cent } \delta i = |\Delta i|_{max}/i(\tau) = 0.37 A_i \tau_0 / A_i \tau$$

$$= 0.37 (\tau_0/\tau) \times 100 \quad (11.12)$$

This error can be minimized by boosting the control current with an initial compensating pulse of amplitude I_0 and duration τ_{com}, sufficient to produce the charge Q_0.

A substantially overdamped response ($\Delta > 0$, $\gamma > 1$) is achieved when $R_{eq} \ll R_{eq,\,cr}$. Retaining only the first two terms in the power expansion of the second term in the parentheses of Eq. (11.7), the roots of the characteristic equation are approximately found to be

$$s_1 = -1/\tau_1, \quad s_2 = -1/\tau_2 \quad (11.13)$$

where $\tau_1 \approx L/R_{eq}$ and $\tau_2 \approx CR_{eq}$, such that $\tau_1 \gg \tau_2$.

With an overdamped response, the transients in the deflection system have two components, a "slow" exponential term with a time constant τ_1, and a "fast" component with a time constant τ_2. If we choose the control-current step B_0 so that

$$B_0 = A_0 \tau_1 \quad (11.14)$$

we can fully balance out the "slow" component. Then the deflecting current will be

$$i = A_0 \{t - \tau_2 [1 - \exp(-t/\tau_2)]\} \quad (11.15)$$

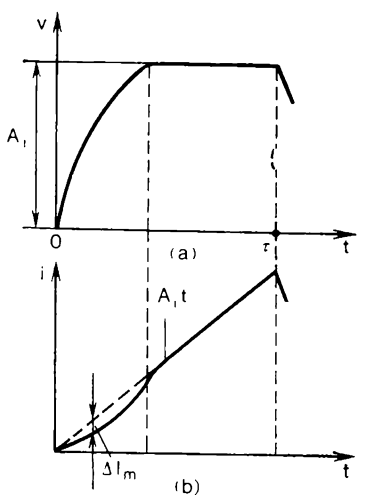


Fig. 11.2

After the transients represented by the exponential term in Eq. (11.15) have died out, the sweep waveform will be delayed by a time τ_2. This time and also the initial nonlinear portion of the sweep might be minimized by reducing the value of R_{sh}. This would, however, lead to a substantial rise in the control current owing to the current drawn by R_{sh}. Because of this, overdamping should

preferably be utilized in sweep generators using a large amount of feedback provided by the voltage across the deflection system, as this permits a reduction in the equivalent internal resistance of the output stage by dropping the low-value resistor R_{sh}. Another disadvantage associated with overdamping is that it takes a long time for the current to return to its initial value at the end of flyback because of the "slow" component, so the speed of the sweep generator is rather low. To avoid this, the negative feedback path should be opened for the duration of flyback and the deflection system should be shunted by a resistor, R_{sh}, whose value is equal to that under critical damping, R_{cr}.

Now we shall take a look at the transient response of the deflection system driven by a staircase control current, $i_0(t) = B_0$. This form of response is shown by a pulsed sweep and also during the flyback of a linear sweep. On substituting $i_0(s) = B_0/s$ in Eqs. (11.5), we obtain the following relations for the three forms of transient response in terms of deflecting current and voltage.

The overdamped response for which $\gamma > 1$, $\tau_1 \gg \tau_2$, and $t > 3\tau_2 - 4\tau_2$:

$$\begin{aligned} i(t) &= i(0) - (R_{eq}/L)[i(0) - B_0][\tau_1 \exp(-t/\tau_1) \\ &\quad - \tau_2 \exp(-t/\tau_2) - (\tau_1 - \tau_2)] \\ &\approx B_0 + [i(0) - B_0] \exp(-t/\tau_1) \end{aligned} \tag{11.16a}$$

$$\begin{aligned} v(t) &= R_{eq}[i(0) - B_0][-\exp(-t/\tau_1) + \exp(-t/\tau_2)] \\ &\approx -R_{eq}[i(0) - B_0] \exp(-t/\tau_1) \end{aligned} \tag{11.16b}$$

—the critically damped response for which $\gamma = 1$:

$$i(t) = B_0 + [i(0) - B_0][1 + (t/\tau_0) \exp(-t/\tau_0)] \tag{11.17a}$$

$$v(t) = 2R_{eq}[i(0) - B_0](t/\tau_0) \exp(-t/\tau_0)] \tag{11.17b}$$

—the underdamped (oscillatory) response for which $\gamma < 1$:

$$i(t) = B_0 + (2R_{eq}/\rho)[i(0) - B_0]\left(\frac{\gamma^2}{\sqrt{1-\gamma^2}} \sin \omega_0 t \sqrt{1-\gamma^2} + \gamma \cos \omega_0 t \sqrt{1-\gamma^2}\right) \exp(-\gamma\omega_0 t) \tag{11.18a}$$

$$v(t) = 2R_{eq}[i(0) - B_0]\left(-\frac{\gamma}{\sqrt{1-\gamma^2}} \sin \omega_0 t \sqrt{1-\gamma^2}\right) \exp(-\gamma\omega_0 t) \tag{11.18b}$$

where $\omega_0 = 1/\sqrt{LC}$.

The settling time of the deflecting current in a pulsed sweep and the flyback time of a linear sweep are maximal for the overdamped response.

If the generator stage operates within the linear portion of the characteristics, the deflection system is shunted by R_{sh} whose value

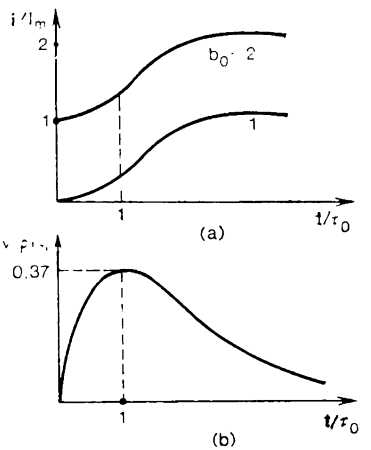


Fig. 11.3

is equal to or somewhat higher than R_{cr}. The functions defined by Eqs. (11.17) are shown in Fig. 11.3*a* and *b*, where $I_m = B_0 - i\,(0)$, and $b_0 = B_0/I_m$.

Referring to Fig. 11.3*b*, $v/\rho I_m$ is a maximum at $t = \tau_0$ and is equal to 0.37. Because the deflection system has a fairly high characteristic impedance (30 kΩ in vacuum-tube sweep generators and 150Ω in transistor sweep generators), a change of 0.1 to 10 A in the control current may cause the voltage across the deflection system to run into hundreds or tens of volts, respectively, thereby moving beyond the linear portion of the characteristics of the vacuum tubes or transistors.

In the underdamped (oscillatory) case, the rise time τ_r of the deflecting current decreases, and the output waveform is a train of waves progressively decreasing in amplitude (known as *ringing*). The relative rise time $\omega_0\tau_r$ of the deflecting current and the relative overshoot amplitude δI_m depend on γ which defines the degree of damping provided by the resonant circuit. Sometimes, the settling time of $i\,(t)$ is evaluated in terms of the damping factor k_d for the envelope of sinewaves [33]:

$$k_d = \exp\,(-\gamma\omega_0\tau_d) \qquad (11.19)$$

where τ_d is the damping time, or the time it takes the waves to die out completely.

Theoretically, the damping time must be a minimum in the critically damped case for which $\gamma = 1$. For practical visual displays, however, it is often convenient to deem that all transients have died out when the luminous spot on the CRT screen has practically come to a stop, which occurs when the spot has deflected for not more than a half-diameter. In present-day intensity-modulated CRTs, the spot diameter is 1/500th of the screen radius.

Thus, the spot practically cease moving at $k_d = 2 \times 10^{-3}$. If the deflection system resonant circuit be shunted by R_{eq} whose value somewhat exceeds the critical value, the response will be slightly underdamped (slightly oscillatory), and $\tau_{d,\,min} = 5.8\sqrt{LC}$ will occur at $\gamma = 0.885$ for which the relative amplitude of the first cycle is 2×10^{-3}. For comparison, it may be noted that for the critical case ($\gamma = 1$), $\tau_{d,\,min} = 8.2\sqrt{LC}$. As is seen, even a small decrease in γ leads to a reduction of 40% in the damping time.

11.3. NEGATIVE-CURRENT-FEEDBACK SWEEP GENERATORS

In practical sweep generators of this type, the control current is that of the final stage built around sufficiently high-power tubes or transistors.

As already noted, one way to reduce nonlinear distortion in, and the drift of, the deflection current caused by the nonlinearity and instability of the final stage, is to apply a sufficient amount of negative feedback to the latter.

The load resistors supplying the negative feedback voltage are chosen to be very small so as to avoid a fall in the operating voltage of the deflecting stage which usually draws a heavy current (the resistance values are usually several ohms for transistor final stages and tens or even hundreds of ohms for vacuum-tube stages). To obtain the desired amount of negative feedback in the circumstances, it is usual to place a preamplifier between the sweep generator and the deflection stages.

In a vacuum-tube sweep generator incorporating an ideal preamplifier with an unlimited passband and a constant gain, the critical response would be obtained in practically the same manner as in the isolated resonant circuit examined above. This is because the negative-voltage feedback would raise the equivalent output resistance of the tube, which even without negative feedback markedly exceeds the critical value of the shunting resistor.

In transistor sweep generators incorporating a preamplifier, the transient response is somewhat different. This is because the deflection coils have a small number of turns, so the distributed interturn capacitance of the deflection system and also its ohmic resistance have a lesser effect on the response than the transit-time parameters of the final-stage transistor, namely the collector-junction capacitance C_c and the time constant τ_β^* of collector current. Taken together, this serves to complicate the structure of the automatic control system formed as a result of applying negative feedback to the transistor sweep generator. The transient response of this system will depend on the transit-time parameters of the system itself and also on the place of connection of the deflection system and the load resistor R_0 across which the feedback voltage is developed. In practice, the deflection system and R_0 can be connected in any one of three ways (Fig. 11.4). In the circuit of Fig. 11.4*a*, the negative feedback circuit stabilizes all of the control current, i_C, whereas in the circuits of Fig. 11.4*b* and *c* it controls the deflecting current i and the current flowing in the shunt.

Consider the conditions required for the response of the above circuits to be critically damped. To begin with, we shall make several simplifying assumptions, usually valid in practice. The point is that the preamplifiers most often used in sweep generators are IC

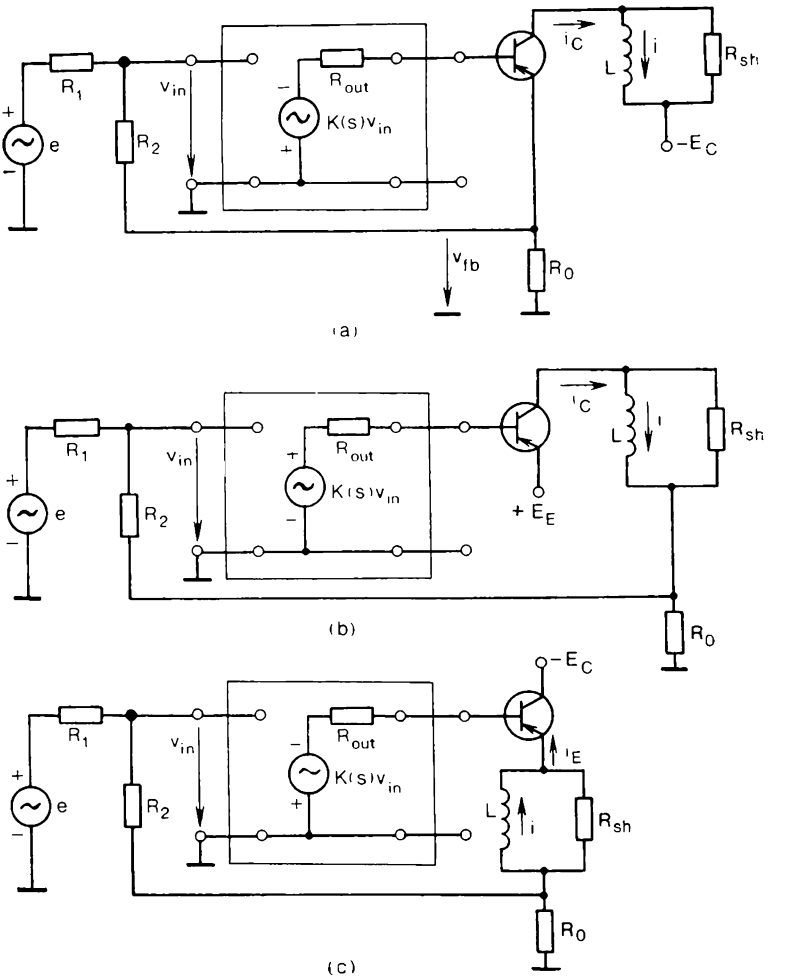


Fig. 11.4

differential opamps with a high input resistance. Also, to avoid parasitic oscillations, the preamplifier passband is reduced by placing a capacitor across the load resistor. So, we may deem that R_{in} is infinite, and approximate the transfer gain by a single-pole rational function

$$K(s) = K_v/(1 + s\tau_{amp}) \tag{11.20}$$

where $\tau_{amp} = (2\pi\Delta f_0)^{-1}$ is the time constant corresponding to the 3-dB passband Δf_0 of the preamplifier.

For r.f. power transistors used in the output stage, such as the Soviet-made 2T803 with $\tau_B \approx 6$ ns, it may be taken that their transit-time behaviour is solely decided by the base junction capacitance and that $\tau_\beta = \beta\tau_B = 0$. Also, recalling that R_0 is very small, the signal at the input to the preamplifier may be written

$$v_{in} \approx \gamma_R e - (1 - \gamma_R)\, v_{FB} \tag{11.21}$$

where

$$\gamma_R = R_2/(R_1 + R_2)$$

$$v_{FB} \approx i_{out} R_0$$

such that $i_{out} = i_C$ for the circuit of Fig. 11.4*b* and $i_{out} = i_E$ for the circuits of Fig. 11.4*a* and *c*.

Sweep generator with a stabilized control current. An a.c. equivalent circuit of a sweep generator with a stabilized control current (Fig. 11.4*a*), drawn up under the above assumptions, is shown in Fig. 11.5. From its analysis [33], it may be concluded that the amount of negative feedback required to minimize the effect of the parameters of the active elements on the deflecting current $i_\sim$, is subject to the following inequalities:

$$K \gg 1, \quad KR_0 \gg (r_{in} + R_B)/\beta \tag{11.22}$$

where $R_B = R_{out} + r'_B$ and $K = K_v (1 - \gamma_R)$.

Subject to the inequalities (11.22), the Laplace transform of the deflecting current, $i_\sim (s)$, in the case of a step input $e = E/s$ may be written

$$i_\sim (s) = \frac{[\alpha - sC_c (r_E + R_0)]\, I}{(a_3 s^3 + a_2 s^2 + a_1 s + 1)\, s} \tag{11.23}$$

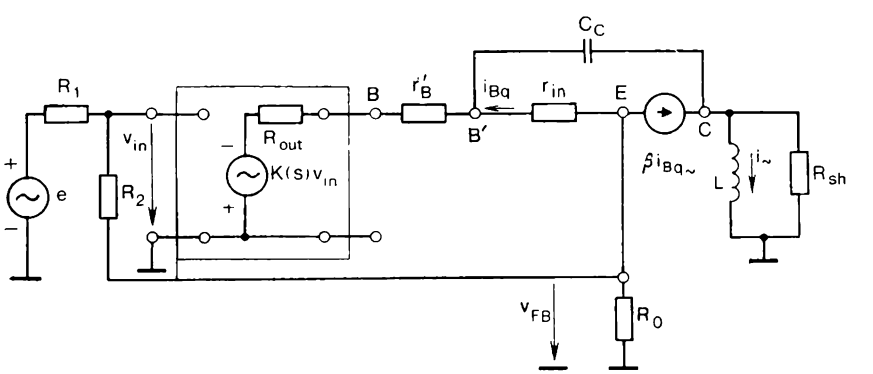


Fig. 11.5

where

$$I \approx ER_2/R_0R_1$$

$$r_E = r_{in}/(1 + \beta)$$

$$a_3 = \tau_{amp}C_c\left[\frac{L(r_E + R_0 + R_B}{KR_0} + \frac{\tau_{ds}R_B(r_E + R_0)}{KR_0}\right]$$

$$a_2 = LC_c + (1 + r_E/R_0)(\tau_{ds}\tau_{out} + \tau_{ds}\tau_{amp} + \tau_{amp}\tau_{out})/K$$

$$a_1 = \tau_{ds} + (1 + r_E/R_0)(\tau_{out} + \tau_{amp})/K$$

$$\tau_{ds} = L/R_{sh}$$

$$\tau_{out} = C_cR_B$$

where τ_{ds} is the time constant of the deflection system.

As is seen from Eq. (11.23), if the inequalities (11.22) are well satisfied, the coefficient a_3 becomes very small, so that its effect will only be felt within the initial portion of the transient response. Therefore, the conditions required to minimize the free response time can be found by using an abbreviated form of the characteristic equation [the denominator of Eq. (11.23)]:

$$a_2s^2 + a_1s + 1 = 0 \tag{11.24}$$

which has the following roots:

$$s_{1,2} = -\frac{1}{2}(a_1/a_2)\left(1 \pm \sqrt{1 - 4a_2/a_1^2}\right)$$

Subject to the equality

$$a_1^2 = 4a_2 \tag{11.25}$$

the critically damped response can be obtained, if we take $R_{sh} = R_{cr}$. Using Eq. (11.25) we obtain that R_{cr} and τ_0 will respectively be

$$R_{cr} = \bar{R}_{cr}2\rho \tag{11.26a}$$

$$\tau_0 = \frac{\tau_{out} + \tau_{amp}}{K} + \sqrt{LC_c + \frac{\tau_{amp}\tau_{out}}{K}} \tag{11.26b}$$

where

$$\bar{R}_{cr} = \left(\frac{\tau_{out} + \tau_{amp}}{2K\sqrt{LC_c}} + \sqrt{1 + \frac{\tau_{amp}\tau_{out}}{KLC_c}}\right)^{-1}.$$

and $\rho = \sqrt{L/C_c}$.

Equations (11.26) are developed under the practically sound assumption that r_E/R_0 is usually much smaller than unity. From Eqs. (11.26) and Fig. 11.6 it follows that the time constant τ_0 of the critically damped system is longer than that of an isolated resonant circuit (see Fig. 11.1) because of the limited passband of the preamplifier and the presence of a resistive component in the capacitance arm of the equivalent circuit of the deflection system. An

increase in the preamplifier gain which controls the amount of negative feedback, minimizes the effect of these parameters, and the difference in dynamic behaviour between the circuits in Figs. 11.1 and 11.4*a* progressively decreases.

Equations (11.26) will yield results satisfactory for practical purposes subject to the following inequality [33]

$$\Delta f_0 > 10\,(K\,2\pi\sqrt{LC_c})^{-1} \tag{11.27}$$

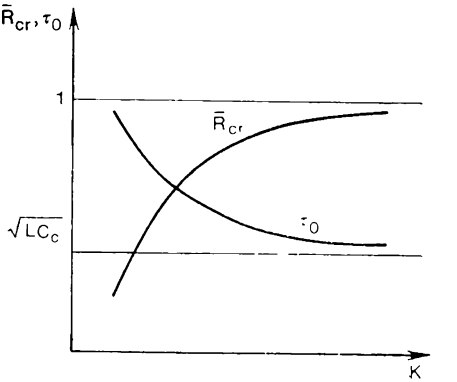


Fig. 11.6

Sweep generator with deflecting current stabilization over the emitter circuit. An equivalent circuit of such a sweep generator (see Fig. 11.4*c*) appears in Fig. 11.7. Subject to the inequalities (11.22), the coefficients of the abbreviated characteristic equation applicable to this case are

$$a_2 \approx L\,(\tau_{out} + \tau_{amp})/KR_0$$

$$a_1 \approx \tau_{ds} + \tau_{amp}\left[1 + \frac{R_B/(1+\beta) + r_E}{R_0}\right]\Big/ K + \tau_{out}\,(1 + r_E/R_0)/K \tag{11.28}$$

Using Eq. (11.25), we get

$$\bar{R}_{cr} = K\left[-0.5\,(\rho/R_0 + \bar{\tau}_{amp} + \bar{\tau}_{out}) + \sqrt{R_B K\,(1 + \tau_{amp}/\tau_{out})/R_0}\right]^{-1} \tag{11.29a}$$

$$\tau_0 = \sqrt{LC_c R_B\,(1 + \tau_{amp}/\tau_{out})/KR_0} \tag{11.29b}$$

where

$$\bar{\tau}_{amp} = \tau_{amp}/\sqrt{LC_c}$$

$$\bar{\tau}_{out} = R_B/\rho$$

Plots of the functions $\bar{R}_{cr}(K)$ and $\tau_0(K)$ appear in Fig. 11.8, where it is assumed that $\tau_{amp}/\tau_{out} \gg 1$ and L/R_0 is markedly in

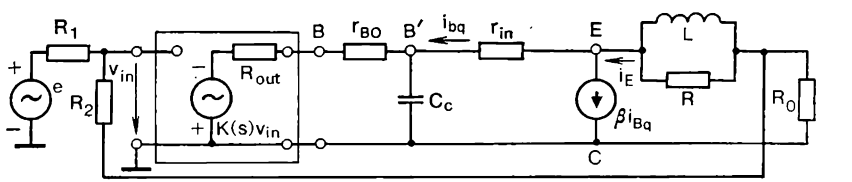


Fig. 11.7

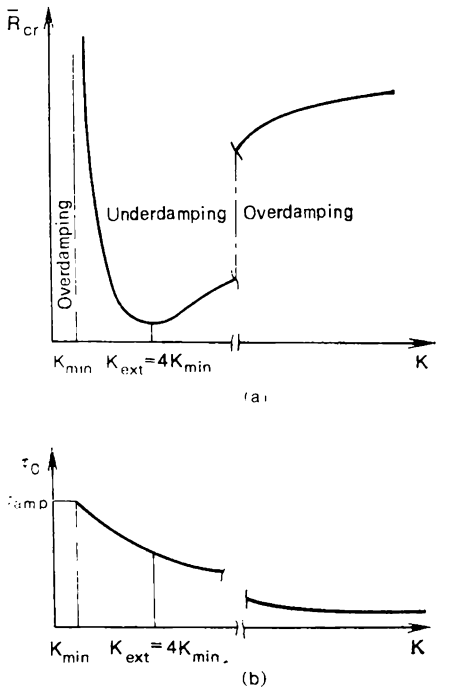


Fig. 11.8

excess of τ_{amp} and τ_{out}, which is often the case in practice. From Eq. (11.29) it follows that in the sweep generator of Fig. 11.4*c* the speed increases with an increase in the passband of the preamplifier and a decrease in the effective inductance of the equivalent resonant circuit of the deflection system due to the effect of negative feedback. Free oscillations will be critically damped at high values of the gain, $K \geqslant K_{min}$, where

$$K_{min} \approx L/4R_0\tau_{amp} = \pi\Delta f L/2R_0 \qquad (11.30a)$$

The time constant corresponding to K_{min} is

$$\tau_{0,max} = (\pi\Delta f_0)^{-1} \qquad (11.30b)$$

In practice, it is often advisable to drop R_{sh}. Then, if $K > K_{min}$ to a first approximation, the time constant will remain the same as follows from Eq. (11.30b). If $K < K_{min}$, the free oscillations will be overdamped, with the "slow" component having a time constant $\tau_1 \approx L/KR_0$ and the "fast" component having a time constant $\tau_2 \approx \tau_{amp}$. Therefore, the speed of the sweep generator in such a case will be determined by the value of τ_1. For $\tau_{amp}/\tau_{out} \gg 1$, we get

$$\tau_0 \approx \sqrt{L\tau_{amp}/KR_0} \approx \sqrt{L/2\pi R_0\Delta f_0 K} \qquad (11.31)$$

It is seen from the above relation that the speed of the sweep generator depends on the time constant of the deflection system and the passband of the preamplifier, rather than on the value of C_c.

Simulation of the circuit in Fig. 11.4*c*, with the inclusion of the cubic term in the characteristic polynomial, has shown [33] that to a first approximation it is legitimate to take $\bar{R}_{cr}$ as defined by Eq. (11.29a). The minimum transient time, τ_{min}, can be estimated accurate to within 10% by the equation

$$\tau_{min} = 3.6\tau_0 \qquad (11.32)$$

Comparison of sweep generators in terms of speed. The speed of the sweep generator in Fig. 11.4*a* depends on the preamplifier passband Δf_0 very little. In fact, it is solely decided by L and C_c over a wide range of gain values ($K = 100$ to 500). As regards the circuit of Fig. 11.4*c*, its speed is controlled by K and Δf_0 nearly to the same extent.

If we take the same values of L and C_c for both circuit configurations, that in Fig. 11.4*c* will be more fast-acting over the above range of gain values if the preamplifier has a sufficiently broad passband ($\Delta f_0 = 100$ to 200 kHz). If, however, the preamplifier has a relatively narrow passband (less than 10 kHz), the more fast-acting circuit will be that in Fig. 11.4*a*. The circuit in Fig. 11.4*b* stands midway between the two. Similarly to that in Fig. 11.4*a*, it is fairly insensitive to gain value, but like the circuit in Fig. 11.4*c*, its speed substantially depends on the preamplifier passband. To be more specific, if the preamplifier passband is from 100 to 200 kHz or more, it is inferior in terms of speed to the circuit in Fig. 11.4*c*, whereas at low values of Δf_0 it is inferior to the circuit in Fig. 11.4*a*.

Large-signal speed. So far we have been concerned with the transient responses of sweep generators operating on relatively small signals, such as the linear timebases of radar scopes or the sign channel of alphameric displays where relatively small step changes in the deflecting current are involved. In the coordinate channel of alphameric displays, such responses occur only within some portions of the deflecting current waveform. Spikes on the leading and trailing edges of current pulse or on the return stroke of a sawtooth current will drive the tube or the transistor in the output stage to the overdriven (critical) or saturation region, unless suitable measures have been taken to the contrary. If this is allowed to happen, the transient response becomes heavily overdamped [33], and the transient response time increases considerably.

To keep the output stage in the active region in the presence of large signals, it is usual to employ either an additional voltage source (*voltage boost*) or a current source having a high differential resistance (*current boost*). Voltage boosting is based on an impulse increase in the supply voltage of the output stage over a time interval equal to the total transient time or a part thereof, so as to avoid completely or in part the fall in voltage across the amplifying element to its critical value. The magnitude of the voltage boost pulse should be chosen such that the decrease in the voltage across the amplifying element owing to a spike of amplitude V_m will leave the operating point within the active region of the characteristics.

In this way, it is possible to keep the transient time to a minimum by eliminating the time interval over which the resonant circuit of the deflection system is heavily overdamped. Then the transient response time will be governed by the same factors as in the small-

signal case. Power dissipation will not practically increase, because the voltage boost is effective over a short time interval proportional to the transient time.

The method outlined above may be used only if the spike amplitude remains below the maximum rated collector voltage of the output transistor. Otherwise, the transistor might be damaged even by short but large spikes if there is a mismatch in waveform between the voltage spikes on the deflection system and the voltage boost pulse. To avoid this, the boost voltage must be limited to the maximum rated collector voltage of the transistor (50 to 400 V). In practice, the above method can be employed only in rare exceptional cases, involving high-voltage transistors and very heavy deflecting currents. The spike amplitude can be brought down by a decrease in the characteristic resistance of the equivalent resonant circuit of the deflection system and an increase in C_{out}. Most transistor sweep generators have to be content with only a partial boost of the fall in the supply voltage. The amplifying element is driven into the saturation region, but its stay there is markedly shorter as compared with that when no voltage boost is used.

With current boost, an element with a high differential resistance exceeding the critical resistance of the deflection-system resonant circuit is switched in between the power supply and the deflection system for the duration of transients. The d.c. resistance of this element is chosen to be small so as to reduce the drain on the power supply. This element may advantageously be a choke whose inductance exceeds that of the deflection coils by at least an order of magnitude, and whose current is somewhat greater than the maximum deflecting current. The choke will, to a first approximation, act as a current source with an infinitely high internal resistance for the duration of transients in the deflection system. The saturated transistor in the output stage will not shunt the deflection system, so the deflecting current may readily be critically damped.

When a choke is placed in series with the deflection system and the saturated transistor, the deflection-system current tends to rise to the value of the choke current during the rise time of the input signal. When the current flowing through the saturated transistor becomes equal to βi_{Bq}, the transistor moves out of the saturated region, and the deflection-system current becomes nearly equal to its steady-state value. Because at that instant the choke current usually exceeds the current in the deflection system and output transistor, the ensuing transients due to the difference in value between these currents may be avoided, using various circuit arrangements. Their choice depends on the type of sweep generator used [33].

11.4. OPAMP INTEGRATOR SWEEP GENERATORS

As already shown, if the electron beam is to be swept across the screen in a linear fashion, the sweep voltage must be a ramp (or sawtooth) waveform supplied by a ramp voltage generator ordinarily containing an amplifier to which a large amount of negative feedback is applied. In some cases, the discharge current of the capacitor in the ramp voltage generator may be stabilized by the preamplifier of a current timebase generator. In this way, a ramp (sawtooth) voltage generator and a ramp (sawtooth) current generator will be built as a single unit. This combination is usually called an opamp-integrator sawtooth current generator.

Most often, an opamp-integrator sawtooth current generator is built in any one of two configurations. In the configuration of Fig. 11.9*a*, the deflection system is not included in the feedback path, the current-stabilizing opamp is loaded into the base-emitter circuit of the output stage, and the feedback voltage is taken off R_0 placed in the emitter lead of the same stage.

In the configuration of Fig. 11.9*b*, the current-stabilizing opamp is loaded into the collector circuit of the output stage, and the feedback voltage is taken off the deflection system. Because transistor-driven deflection circuits have a low resistance, it is usual to enhance the performance of the sawtooth current generator in Fig. 11.9*b*

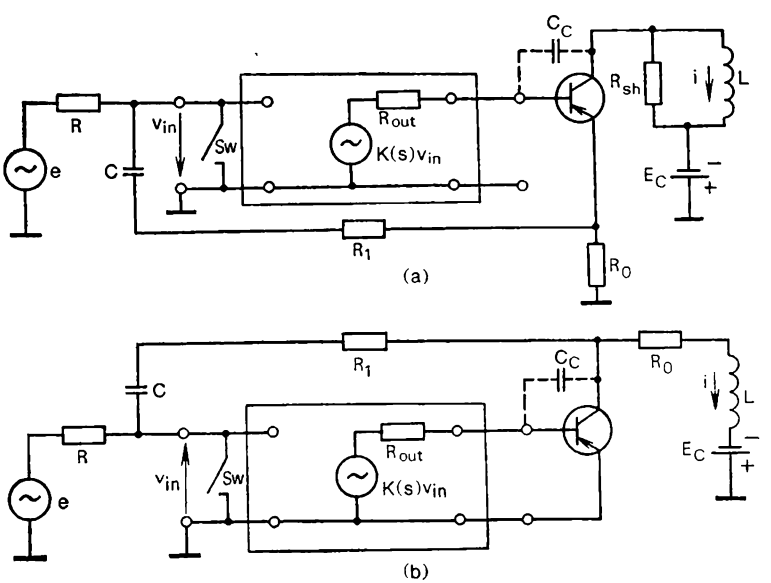


Fig. 11.9

by connecting the deflection system in series with a resistor, R_0, serving as load for the opamp.

The two circuits in Fig. 11.9 operate by about the same principle. When the switch Sw is opened, the capacitor C begins to be charged by a nearly constant current from a constant-voltage source $e = E$, as is done in the Miller integrator ramp voltage generators examined in Chap. 9. The resultant voltage across the negative feedback circuit (which is resistive-capacitive in its effect) will be trapezoidal, that is, it will have both a step portion and a ramp portion. Thus, in the generator of Fig. 11.9*a*, negative feedback will stabilize the trapezoidal current of the output stage, and in the generator of Fig. 11.9*b*, the trapezoidal voltage at the collector. If we deem the current in the negative feedback circuit constant, $I_0 \approx E/R_s$, and neglect the transit-time effects in the preamplifier (τ_{amp}) and the output stage (τ_C), then, in order to obtain a deflecting current rising linearty with time, $i = A_i t$, the circuit parameters must be chosen such that

$$L/R_{sh} = CR_1 \tag{11.33a}$$

for the circuit in Fig. 11.9*a*, and

$$L/R_0 = CR_1 \tag{11.33b}$$

for the circuit of Fig. 11.9*b*. Then the rate of rise of the deflecting current will be

$$A_i = E/CR_sR_0 \tag{11.34}$$

In practice, because the preamplifier (most often, an IC differential opamp) has a finite gain, the current in the negative feedback circuit does not remain constant. This causes the deflecting current to show a degree of nonlinearity which is further aggravated by failure to satisfy exactly the conditions defined by Eqs. (11.33). The transit-time parameters, which tend to limit the preamplifier passband, bring down the rate of change A_i of the deflecting current and produce a further nonlinearity within the initial portion of the waveform. Analysis [33] shows that the effect of these parameters can be minimized by increasing the amount of negative feedback which is in turn dependent on the gain of the preamplifier.

The charge on the integrating capacitor is restored after the control pulse ceases. Thus, the speed of response of the sweep generator in question depends not only on the time it takes the deflecting current to attain its steady-state value, but also on the time required for the capacitor C to recover its initial charge. If the capacitor charge fails to be restored by the time when the next control pulse arrives, the sweep speed A_i will change. Therefore, sweep generators built along the lines illustrated in Fig. 11.9 should include appropriate circuits to speed up the recovery of the charge on the integrating capacitor.

Because the negative feedback circuit of the generator is now open, the magnetic energy stored by the deflection system should preferably be dissipated by switching in an appropriate damping resistor to obtain critical damping.

11.5. SAWTOOTH (RAMP) CURRENT GENERATORS FOR TELEVISION

The sawtooth (ramp) current generators used in the vertical (frame) and horizontal (line) scanning circuits of TV receivers are designed to meet requirements differing markedly from those to be satisfied by the sweep generators employed in radar indicators. The typical TV sweep is continuous, with a relatively short retrace (flyback) time accounting for 5% and 10% of the sweep period for the frame and line scanning circuits, respectively. The requirements for linearity and stability are usually fairly lenient ($\xi_l = 5$ to 10%), with principal emphasis placed on simplicity of control, economy, small size, and low cost.

Because the duration of the vertical sweep is hundreds of times that of the horizontal sweep, the respective sweep generators differ in a substantial way.

Vertical scanning. Because the rate of vertical (frame) scanning is relatively low, the distortion in the waveform of the frame deflecting current caused by the stray capacitance are ordinarily negligible, so, in a first approximation, this capacitance may be ignored in defining the waveform of the scan portion. This is why this capacitance is omitted from the equivalent circuit in Fig. 11.1. On the other hand, in the case of a continuous sweep (Fig. 11.10) the deflection system carries a reverse current, $I_0 \approx 0.5I_m$, when the scan portion just begins. Accordingly, the Laplace transform of the deflecting current over the time interval τ may be written

$$i(s) = \gamma_R \frac{i_0(s) - LI_0/R_{eq}}{1 + s\tau_{ds}} \tag{11.35}$$

where $\gamma_R = R_{eq}/(R_{eq} + r)$ and $\tau_{ds} = L/(R_{eq} + r)$.

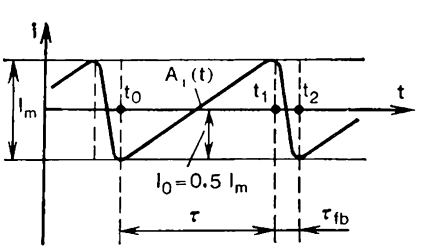


Fig. 11.10

In the case of a trapezoidal control current,

$$i_0(s) = (A_0/s + B_0)/s$$

the inverse Laplace transform of Eq. (11.35) will be

$$i(t) = A_i t - I_0 + (I_0 + A_i B_0/A_0 - \tau_{ds} A_i)[1 - \exp(-t/\tau_{ds})] \quad (11.36)$$

where $A_i = \gamma_R A_0$.

For $i(t)$ to vary linearly with time, the step B_0 must be such that

$$B_0 = A_0(\tau_{ds} A_i - I_0) A_i \quad (11.37a)$$

Since $A_i = I_m/\tau$, we get

$$B_0 = I_m(1 + r/R_{eq})(\tau_{ds}/\tau - 0.5) \quad (11.37b)$$

In practice, variations in the parameters of the tube or transistor used in the output stage usually result in that the condition defined by Eq. (11.37) is not satisfied in full, and the deflecting current becomes nonlinear. The degree of nonlinearity increases with that of the V-A characteristic of the tube or transistor and also owing to the magnetizing current which flows in any transformer-coupled circuit. In either case, the waveform of the deflecting current can be improved by using frequency-dependent negative feedback and including compensating RC-networks. During vertical retrace, the tube or transistor in the output stage is driven to cutoff, and the damping resistor ensures that the deflection-system resonant circuit operates under slightly underdamped conditions.

Horizontal scanning. In this case, the basic problems are to minimize the flyback time (to not over a few microseconds) and to reduce power dissipation. Calculations [41] show that in practice, assuming that all of the magnetic energy stored by the deflection system is dissipated during retrace, power consumption is over 20 W. Because of this, instead of damping resistors practical horizontal sweep generators use active damping elements, such as diodes or transistors, which turn off for the duration of retrace. If so, the retrace time is determined by a half-cycle of oscillation in the resonant circuit of the deflection system. It is stabilized by trimming capacitors connected in parallel with the deflection system. Therefore,

$$\tau_{fb} \approx \pi\sqrt{LC} \quad (11.38)$$

Now the magnetic energy stored by the deflection system by the end of retrace is utilized to nearly double the deflecting current during scan. This is illustrated in Fig. 11.11*a* where the switch simulating the action of the active elements in the output stage is bidirectional (reciprocal). During the scan interval, $t_0 - t_1$, the switch is closed and the resonant circuit is shunted by $R_{sw} \ll R_{cr}$,

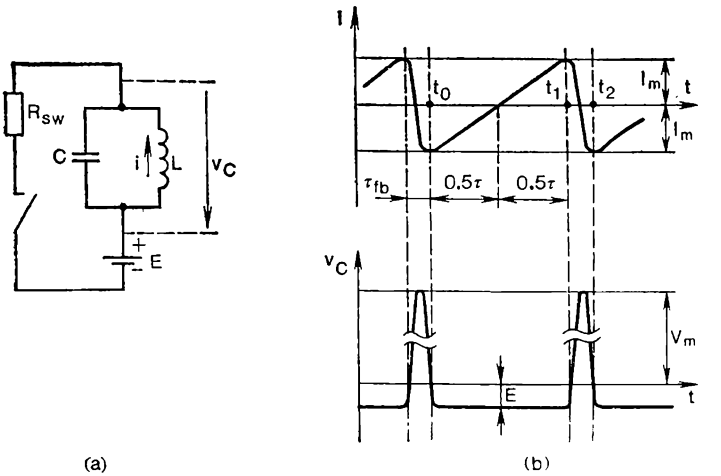


Fig. 11.11

so that the resonant circuit is heavily overdamped. Neglecting C in this case, we may deem that the deflecting current varies exponentially (Fig. 11.11*b*):

$$i\,(t) = E\,[1 - \exp\,(-t/\tau_{sw})]/R_{sw} - I_m \exp\,(-t/\tau_{sw}) \qquad (11.39)$$

where $\tau_{sw} = L/R_{sw}$.

From Eq. (11.39), the initial rate of change, A_i, and the nonlinearity, ξ_i, of the deflecting current are found to be

$$A_i = \frac{E/L}{1 - 0.5\tau/\tau_{sw}} \approx (E/L) \quad \text{for} \quad \tau/\tau_{sw} \ll 2 \qquad (11.40a)$$

$$\xi_i = \tau/\tau_{sw} \qquad (11.40b)$$

During scan, the voltage across the deflection system is $v_{ds} \approx E$. During retrace, $t_1 - t_2$, the value of v_{ds} rises to its peak value

$$V_m \approx I_m \sqrt{L/C} = \pi L I_m/\tau_{fb} \qquad (11.41)$$

which is 100 V and higher in transistor horizontal sweep generators. The circuit of a practical transistor horizontal sweep generator, simplified by omitting minor elements, is shown in Fig. 11.12. During scan, the power transistor in the output stage, containing the deflection system in its collector lead, accepts via a transformer, $Tr1$, a positive pulse of base current, $I_B > I_{B(sat)}$, where $I_{B(sat)} = I_m \beta$. As a result, the scan portion of the sweep is generated while the transistor is at saturation, being reverse-biased during the first half of the scan and forward-biased during the second half, so that the transistor operates as a bidirectional (reciprocal) switch. If the

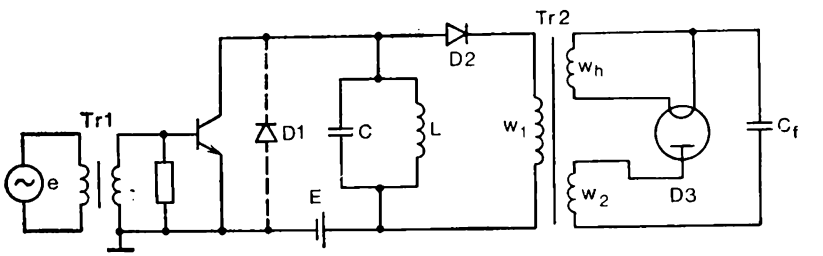


Fig. 11.12

transistor is a conventional power type in which the emitter and collector junctions differ in size, the transistor is shunted by a damping diode shown by the dashed line in Fig. 11.12.

During flyback, the transistor base accepts a negative current pulse of large amplitude so as to turn off the transistor quickly. The cosine-wave voltage pulse of about 100 V amplitude, developed across the deflection coils, is impressed on the primary winding of a high-voltage transformer $Tr2$ with a turns ratio of $n \geqslant 100 : 1$. The filter capacitor C_f and a vacuum diode $D3$ in the secondary winding of the transformer make up a rectifier for the high voltage fed to the second anode of the picture tube.

Instead of a vacuum diode, the H.V. rectifier may use H.V. semiconductor structures, such as selenium stacks which are usually arranged in a voltage doubler or tripler circuit.

If the horizontal output stage is built around a vacuum tube (usually, a beam power tetrode), its circuit is more complicated. Because the tube cannot act as a bidirectional (reciprocal) switch, one is forced to use a damping diode. Also, because the tube has a high output resistance, the deflection system is connected via a step-down transformer. The operation of, and basic relations for, the vacuum-tube line sweep oscillator are examined in [4].

REVIEW QUESTIONS AND PROBLEMS

11.1. Draw the waveform of the control current required to produce a sawtooth (ramp) current in the deflection coils.

11.2. Draw the waveform of the deflecting current in the case of a trapezoidal control current, assuming the critical and overdamped response. Point out the difficulties associated with the overdamped case.

11.3. Draw the waveform of the deflecting current in the case of a stop input (control) current for the underdamped, critically damped and overdamped cases.

11.4. Draw the three types of negative-current-feedback transistor sweep generator you know; explain their differences.

11.5. Draw the equivalent circuit of the sweep generator in Fig. 11.4*a*; derive the functions $R_{cr}(K)$ and $\tau_0(K)$ from analysis of the simplified step-input transient response. Compare the effects of the amplifier's time lag, the transition (junction) capacitance of the output transistor, and the inductance of the deflection coils on the two functions.

11.6. Work out Problem 11.5 for the sweep generator in Fig. 11.4*c*.

11.7. Compare the performance of the sweep generators in Figs. 11.4*a* and *c*.

11.8. Explain the rationale of using voltage and current boost under large-signal conditions; explain how voltage and current boosting works.

11.9. Draw the circuits of the two opamp integrator sweep generators you know; explain how they differ in operation.

11.10. Derive Eqs. (11.33) and (11.34).

11.11. Determine the distortion produced in the deflecting current by failure to meet the conditions (11.33) exactly.

11.12. Draw a simplified equivalent circuit for the transistor vertical scanning output stage of a TV receiver; plot the associated waveforms and determine the relations between the circuit parameters required to obtain a sawtooth (ramp) current.

11.13. Draw a simplified equivalent circuit for the transistor horizontal scanning output stage of a TV receiver; plot the associated waveforms, and explain its operation. Determine the parameters of the scan and retrace portions of the sweep; explain the operation of the circuit in Fig. 11.12.

12 NEGATIVE-RESISTANCE PULSE CIRCUITS

12.1. GENERAL

Using semiconductor devices whose volt-ampere characteristics show a portion with a negative slope, $g_m = di/dv < 0$, it is possible to build regenerative pulse circuits in which positive feedback occurs within the devices, that is without an external feedback circuit. In some cases, this simplifies circuit configuration and serves to improve speed of response, load capacity, etc.

There are basically two types of negative-resistance characteristics that a device might have. These are *N-type* (Fig. 12.1*a*) and *S-type* (Fig. 12.1*b*). In the former, the current is a unique function of voltage, so these devices are voltage-controlled. In the latter, the voltage is a unique function of current, so they are current-controlled. The N-type characteristic is exhibited by tunnel diodes which generate pulses in the nanosecond range of duration. The S-type characteristic is shown by avalanche transistors (when connected in a common-emitter circuit) and four-layer structures (thyristors) which can generate strong pulses of large amplitude.

12.2. TUNNEL-DIODE PULSE CIRCUITS

The basic circuit using a tunnel diode is shown in Fig. 12.2*a*. In the diagram, C_s is the stray capacitance combining the circuit

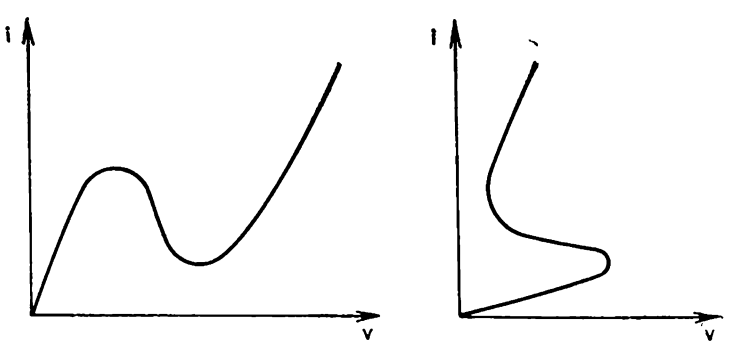


Fig. 12.1

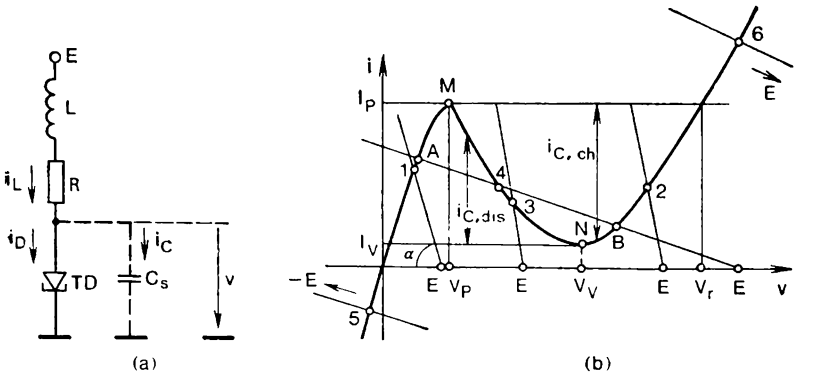


Fig. 12.2

(wiring) capacitance, the load capacitance, and the transition (junction) capacitance which ranges in value from units to tens of microfarads. The tunnelling portion of the characteristics (Fig. 12.2*b*) has a negative slope between points M and N, where the average resistance is negative, $-\bar{\rho}$. The negative-resistance characteristic is defined in terms of the peak-point current I_P, the peak-point voltage V_P, the valley-point voltage V_V, and the voltage range V_r. The typical values of these parameters are: $I_P = 1$ to 50 mA, $V_P = 75$ to 330 mV, $V_V = 300$ to 600 mV, $V_r = 0.4$ to 1.3 V, and $I_P/I_V = 5:1$ to $10:1$.

Depending on the value of supply voltage E and $R = (\tan \alpha)^{-1}$, the circuit in Fig. 12.2*a* may have five equilibrium states illustrated in Fig. 12.2*b* [30]. If the load line intersects one of the rising portions of the volt-ampere characteristic of the tunnel diode and $R < \rho$, a state of stable equilibrium will occur at points *1* and *2*. Now, a deviation of the emitter voltage of the tunnel diode from its mean value owing to some factor will cause the capacitive current to flow in a direction such that the rate of displacement of the operating point will tend to minimize the initial deviation. This case is characteristic of one-shot multivibrators with an initial stable state at point *1* or *2*.

At point *3* where R is likewise less than ρ, the above effect and, as a consequence, a state of stable equilibrium will occur if the inductance is sufficiently small, so that $L/R < \rho C_s$. This condition is typical of tunnel-diode amplifiers. If the above condition is not satisfied, the state of equilibrium at point *3* will be unstable, and the circuit will be running free. Then, if $L/R \gg \rho C_s$, the oscillations will be of relaxation type, such as generated by a free-running multivibrator.

For triggered operation, it is essential that $R > \rho$, whereas L is not a mandatory element. If the supply voltage E is chosen such that the load line crosses all the three portions of the V-I curve (points A, *4*, and B in Fig. 12.2*b*), the circuit will have two states of stable equilibrium (at points A and B) and one state of unstable equilibrium (at point *4*), so it can operate as a bistable multivibrator. If the load line crosses only one of the rising portions of the curve (at point *5* or *6*), the circuit will operate as a Schmitt trigger; it will act as a pulse-shaping circuit if the Q-point is initially positioned outside the hysteresis loop.

The switching time of the tunnel diode from the low-voltage state (point A) to the high-voltage state (point B) and back depends on the time it takes the stray capacitance C_s to charge and discharge. As a rule, the current in the LR load (i_L) has no time to change during this interval, and the capacitive current, i_C, is determined by the ordinates $i_{C,ch}$ and $i_{C,dis}$ lying in the plot of Fig. 12.2*b* between the I_P or I_V levels and the diode characteristic. If we deem, in a first approximation, that the portions MN and NE are symmetrically linear, then the switching time from the low-voltage to the high-voltage state may be written

$$\tau_r \approx 2C_s (V_r - V_P)/I_P (1 - I_V/I_P) \tag{12.1}$$

For practical circuits with a finite load resistance and driven by signals having a finite rise time, the switching time is longer in comparison with what is given by Eq. (12.1), but it depends on the diode parameters in about the same manner.

In the pulse circuits to be discussed later, the switching time is negligible in comparison with the duration of the transients related to changes in the energy stored by the reactances, above all the coils. Also, the speed of cascaded pulse circuits using tunnel diodes is often affected by buffer amplifier stages and is substantially lower than that of the individual stages arranged as shown in Fig. 12.2*a*. Therefore, one of the problems facing the circuit designer is to utilize the inherent speed of tunnel diodes to advantage.

Tunnel-diode multivibrators. The basic circuit of a tunnel-diode free-running multivibrator is shown in Fig. 12.3*a*. The divider R_1/R_2 serves to position the Q-point on the negative slope of the diode characteristic (Fig. 12.3*b*), subject to the conditions

$$R < \bar{\rho} \tag{12.2a}$$

$$V_V > E - I_0 R > V_P \tag{12.2b}$$

where $R = R_1 \| R_2$ and $R_L \gg R$.

Also, the slope of the dynamic load line at switching must be substantially smaller than that of the negative portion of the tunnel-

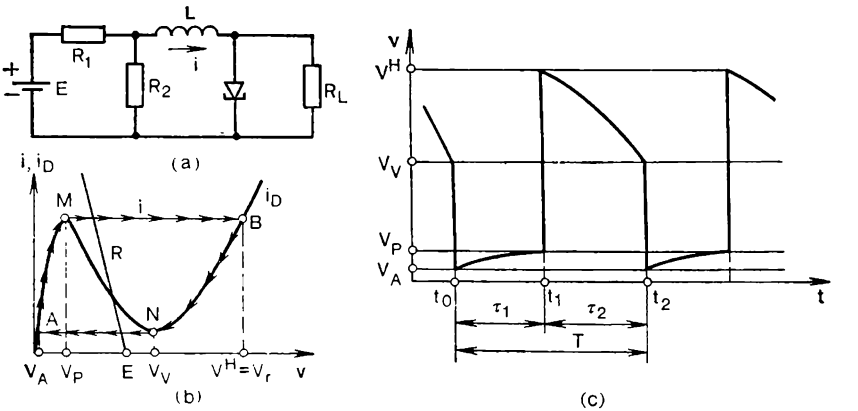


Fig. 12.3

diode characteristic, which is equivalent to requiring that

$$(L + L_s)/R \gg C_s\bar{\rho} \tag{12.2c}$$

where L_s is the stray inductance associated with the diode leads.

To make the waveform of output pulses generated by a tunnel-diode multivibrator as rectangular as practicable, the inductance is chosen to be sufficiently large. Therefore the dynamic load line may be deemed horizontal at switching. The inequalities (12.2) define the conditions for operability of the circuit.

When a tunnel-diode multivibrator is just turned on, the Q-point is positioned within the rising ("tunnelling") portion of the diode characteristic, and the voltage across and the current in the coil and the corresponding diode current i_D begin to rise exponentially. On reaching a maximum at point M (at time t_1, Fig. 12.3*c*), the circuit switches in the forward direction. At switching, the coil current i remains practically unchanged, and the Q-point jumps abruptly to point B. During the interval from t_1 to t_2, the multivibrator remains in a temporary stable state and generates the second half-cycle of oscillation, during which the diode voltage and current decrease exponentially, and the Q-point moves down the "diffusion" part of the characteristic. The circuit switches in the reverse direction at point N (the time t_2), the Q-point jumps abruptly to point A, and within the interval between points A and M the circuit generates the first half-cycle of oscillation from which we began our discussion. Until the circuit switches in the forward direction, the current rises as follows:

$$i = (I_f - I_1)\,(1 - \exp\,[-(t - t_0)/\tau_{01}]) + I_1 \tag{12.3}$$

where $\tau_{01} = L/(R + r_{D,1})$, I_1 is the initial value of current equal to $i\ (t_0) = I_V$ and I_f is the final value of current equal to $E/(R + r_{D,1})$.

On substituting $i\ (t_1) = I_P$, we obtain

$$\tau_1 = \tau_{01} \ln \frac{E/(R+r_{D,1}) - I_V}{E/(R+r_{D,1}) - I_P} \tag{12.4}$$

Until the circuit switches in the reverse direction, the change in the current may be written

$$i = (I_1 + I_f) \exp [-(t - t_1)/\tau_{02}] - I_f \tag{12.5}$$

where

$$\tau_{02} = L/(R + r_{D,2})$$
$$I_1 = i\ (t_1) = I_P$$
$$I_f = (-E + e_{02})/(R + r_{D,2})$$

$r_{D,1}$ and $r_{D,2}$ are the averaged resistances, and e_{02} is the cutoff voltage on the second portion of the curve.

On substituting $i\ (t_2) = I_V$, we get

$$\tau_2 = \tau_{02} \ln \frac{I_P + (e_{02} - E)/(R + r_{D,2})}{I_V + (e_{02} - E)\ (R + r_{D,2})} \tag{12.6}$$

Obviously, $T = \tau_1 + \tau_2$. It follows from Eqs. (12.4) and (12.6) that the stability of T depends on that of the diode parameters and supply voltage. To enhance it, it is desirable to make E as close to e_{02} as practicable. The maximum speed (frequency response) of such circuits is limited by the possibility to satisfy the condition defined by Eq. (12.2) and usually ranges between 10 and 20 MHz.

The circuits of triggered tunnel-diode multivibrators are similar to that of Fig. 12.3*a* with the only difference that the values of R and E are chosen such that the static load line intersects the rising ("tunnelling" or "diffusion") portion of the characteristic. Because of this, the initial stable state gives way to a temporary state under the action of a trigger pulse. The duration of the temporary state is determined, as in the case of a free-running multivibrator, by the decrease or rise of current in the inductance to the limiting values I_V or I_P. After the circuit switches in the reverse direction, the current in a triggered multivibrator is brought back to its initial (static) value. To minimize the recovery time (which usually exceeds the duration of the temporary state), it is usual to include one more pulse or backward diode (the so-called nonlinear bias circuit) [4].

Tunnel-diode flip-flops. The simplest tunnel-diode flip-flop does not differ from the basic multivibrator circuit shown in Fig. 12.3*a*. However, the values of R_1, R_2, and E are chosen sufficiently large for the load line to cross both rising portions of the diode characteristic.

A tunnel-diode flip-flop can operate in any of two modes: current-switching and voltage-switching. In the current-switching mode corresponding to relatively small values of $R = R_1 \parallel R_2$ and $E = E_1$, the change of tunnel-diode current at switchover is a maximum and close to the difference $I_P - I_V$, whereas the change of voltage is relatively small and close to $V_V - V_P$. This mode of operation is usually employed in the case of symmetrical triggering where the initial control signals are unipolar. For reliable operation of such a flip-flop, the initial current value is maintained at switchover by including an inductance. In the voltage-switching mode which is characterized by high values of R and $E = E_2$, the change of output voltage is a maximum, being close to $V_r - V_P$, whereas the change of current is small. In such flip-flops, no inductance is ordinarily included.

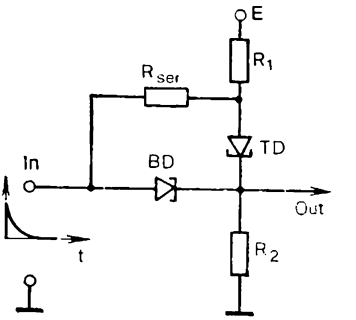


Fig. 12.4

Because the pulses applied to one terminal of a flip-flop must be bipolar, symmetrical triggering calls for the use of suitable circuits to derive such pulses from the original unipolar pulses. In the flip-flop shown in Fig. 12.4, switchover is brought about by unipolar pulses. When the circuit changes state in the forward direction, the backward diode, *BD*, is held at cutoff by the high level of output voltage, and the positive trigger pulse goes to the anode of the tunnel diode via a low-value resistor R_{ser}. When the circuit changes state in the backward direction, the backward diode is conducting and passes the trigger pulse to the cathode of the tunnel diode.

The frequency response of tunnel-diode scaling circuits is several hundred megahertz and more.

12.3. AVALANCHE-TRANSISTOR RELAXATION OSCILLATORS

Basically, avalanche transistors do not differ from the conventional bipolar ones. What difference they have in terms of design and manufacture is related to the requirement for better stability and reliability under shock ionization when the collector voltage is rather high.

When an avalanche transistor is connected in a common-emitter circuit (Fig. 12.5*a*), the avalanche-breakdown region of the collector junction usually lies within the limits $V_\beta \leqslant v_{CE} \leqslant V_M$, where V_β is the critical value of v_{CE} at $\beta = \infty$, and V_M is the avalanche

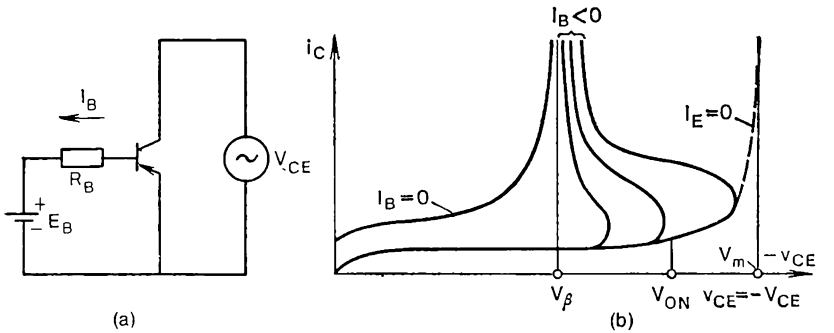


Fig. 12.5

breakdown voltage at which the shock ionization ratio, M, is infinity. Within this region (Fig. 12.5*b*), the volt-ampere characteristic for negative base currents, $I_B < 0$, is S-shaped and has a portion with a negative dynamic resistance, $r_d = -\rho$, where the collector voltage varies between the limits $V_\beta > V_{CE} > -V_{ON}$. The turn-on voltage, V_{ON}, analogous to the firing voltage of thyratrons*, is given [14] by

$$V_{ON} \approx V_M \sqrt[n]{1 + I_{BO}/I_B} \tag{12.7}$$

where $I_B \approx -E_B/R_B$, and n is the exponent whose value depends on the semiconductor material used, type of junction, and the carriers that bring about the avalanche breakdown. For obrupt germanium junctions, $n \approx 3$ if the breakdown is produced by holes, and $n \approx 6$ if the breakdown is caused by electrons. For obrupt silicon junctions, the respective figures are 3.5 and 2 [30].

Among the advantages offered by avalanche-transistor relaxation oscillators are simple circuit configuration, short rise time of the output pulses (not over 0.1 ns), large-amplitude current pulses (up to several amperes) and voltages (as high as 100 V and more), high triggering sensitivity (a few tens of millivolts), a short delay time (0.1 ns and less), good isolation between the input and output circuits, and satisfactory temperature stability. Among the principal disadvantages are a high residual voltage in the ON state (8-10 V and more), and a relatively low pulse repetition frequency (not over tens of megahertz) because of the relatively low average power dissipation of the collector junction [30].

Because of the factors listed above, the best avalanche-transistor relaxation oscillators are multivibrators generating strong nanosecond pulses with a low pulse duty factor, which resemble fairly closely high-power transformerless blocking oscillators.

* Some authors call it "firing voltage" for avalanche transistors as well.—Translator's note.

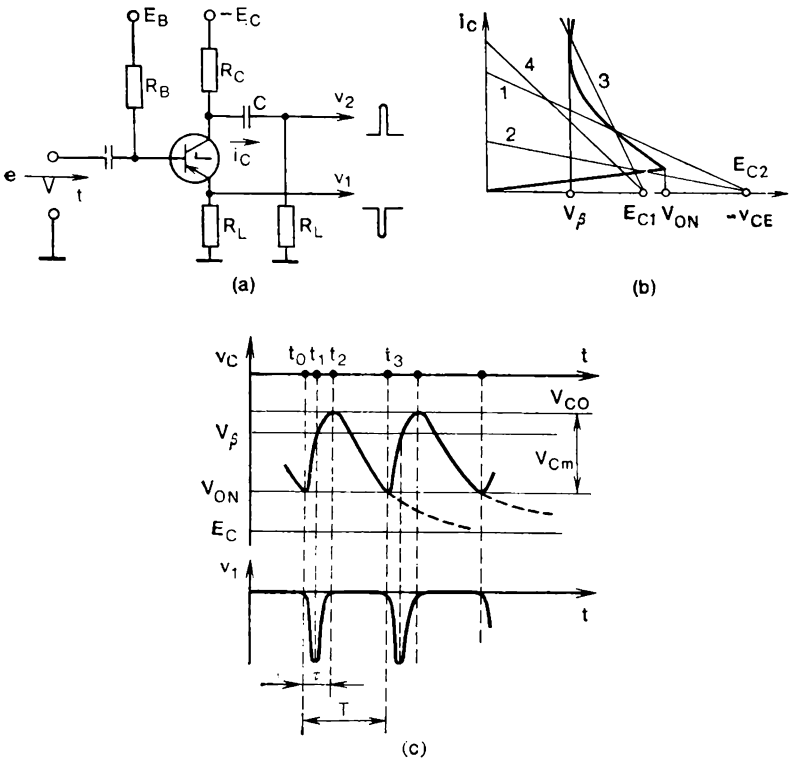


Fig. 12.6

The basic circuit of a simple avalanche-transistor relaxation oscillator is shown in Fig. 12.6*a*. Output pulses in opposite polarity are usually taken off the load resistors R_L. Similarly to tunnel diodes with an N-shaped volt-ampere characteristic, the choice of the value for R_C and $-E_C$ decides whether the relaxation oscillator in Fig. 12.6*a* will operate as a free-running or a one-shot multivibrator (cases *1* and *2* in Fig. 12.6*b*), as a flip-flop (case *3*), or as a Schmitt trigger (case *4*). It is an easy matter to show that the relations between R_C and ρ required for the operation in any one of these modes are reverse of those for tunnel-diode circuits because of the S-shaped static characteristic. More specifically, $R_C > \rho$ for avalanche-transistor multivibrators, and $R_C < \rho$ for a flip-flop.

The operation of an avalanche-transistor relaxation oscillator can be explained in the following manner. As is seen from Fig. 12.6*c*, the time-controlling phase of the preceding oscillation cycle terminates at $t < t_0$, when the transistor is at cutoff and the potential at its collector goes down because the time-controlling capacitor

is being charged. At $t = t_0$, the collector voltage v_C rises to V_{ON} and brings about shock ionization, which is equivalent to closing the positive feedback circuit in a blocking oscillator. The result is that during the interval $t_0 - t_1$ the circuit generates the leading edge of the output voltage v_1 proportional to the discharge current of the capacitor, which is rising during this interval. At t_1, the Q-point moves into the upper (positive-resistance) portion of the volt-ampere characteristic of the avalanche transistor (practically, the resistance may be taken zero at V_β).

At the end of shock ionization (time t_1), the output pulse begins to fall due to the action of the current produced by the removal of the excess minority carriers stored in the emitter junction forward-biased during the previous interval. This process is analogous to the removal of excess minority carriers from the emitter when the switch is turned off by a heavy current (see Chap. 3). At time t_2, all of the excess minority-carrier charge is removed, and this terminates both the discharge of the capacitor and the fall of the output pulse. The duration of the interval $t_1 - t_2$ depends on the excess minority-carrier charge stored in turn determined by the charge on the capacitor C which diminishes over the interval $t_0 - t_1$ and is governed, among other things, by the value of C. Thus, the amplitude and duration of the output pulse in the circuit of Fig. 12.6 depend on the value of C, as is the voltage $-V_{CO}$ to which the time-controlling capacitor discharges.

During the interval $t_2 - t_3$, the capacitor charges and the time-controlling phase of an oscillation cycle is shaped, from which we began our discussion. The duration of this phase can readily be found, assuming $r_{out,\,c\text{-}o} \gg R_C$ and $MI_{CO}\,R_C \ll E_C$:

$$t_3 - t_2 = \tau_{time} \ln\left(1 + \frac{V_{C,m}}{E_C - V_{ON}}\right) \tag{12.8}$$

where $V_{C,\,m} = V_{ON} - V_{CO}$, and $\tau_{time} = CR_C$.

As already noted, it is usual that $\tau \ll t_3 - t_2$, so Eq. (12.8) may be taken as defining the period of oscillation, T. When the relaxation oscillator of Fig. 12.6*a* operates as a one-shot multivibrator, the initial state is stable, because if the load resistor is made sufficiently large, the load line in Fig. 12.6*b* intersects the lower positive-resistance portion of the volt-ampere characteristic. The circuit is started by applying a negative trigger pulse to the base. The trigger pulse reduces the base current in absolute value and this leads to a decrease in V_{ON}.

The delay between the instant when the leading edge of a trigger pulse is applied to the base and the instant when the circuit is actually triggered is caused by the minority-carrier charge stored at the emitter junction. This turn-on delay does not usually exceed a fraction of a nanosecond or several nanoseconds [30]. The remaining

phases (rise, fall and recovery times) are the same as in a free-running multivibrator.

The principal limitation of the simple multivibrators examined above is the instability of pulse repetition frequency and pulse duration. This limitation can be avoided by using crystal-controlled frequency stabilizers [30].

12.4. THYRISTOR RELAXATION OSCILLATORS

Thyristors, or four-layer p-n-structures, are the semiconductor analogues of gas-filled tubes (thyratrons and gas-filled rectifier diodes). The volt-ampere characteristics of a controlled-base thyristor (trinistor) are shown in Fig. 12.7*a*, whereas Fig. 12.7*b* and *c* show, on an enlarged scale, the initial portions of these characteristics for the positive and negative base currents, respectively. It is assumed that $I_B > 0$ is the current entering an n-type base (in a $pnpn$-device) and leaving a p-type base (in an $npnp$-device). For a dinistor (a structure having no base lead), $I_B = 0$.

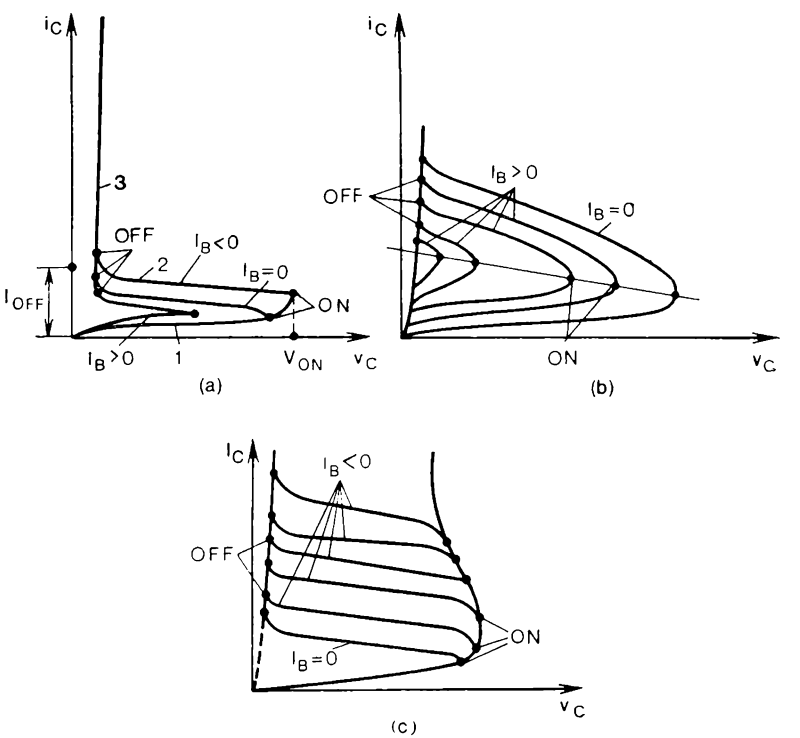


Fig. 12.7

As is seen from Fig. 12.7*a*, thyristors have an S-shaped volt-ampere characteristic, and the dynamic resistance within the interval *2* between the turn-on and turn-off points, ON and OFF, is negative, $r_{d2} = -\rho$, where $\rho = |\, dv_C/di_C \,|$ at $i_B = \text{constant}$. The average values of ρ range from 5-10 to 50-100 kΩ [14]. As in an avalanche transistor, this negative-resistance portion is produced by the shock ionization of carriers at the middle (collector) junction of the device. Within portion *1*, the collector current i_C is very small (units or tens of microamperes) for low-power devices, whereas the resistance is high (units or tens of megohms), so the device may be deemed to be in the OFF state.

Portion *3* is analogous to the volt-ampere characteristic of a conventional crystal diode: the device is in the ON state with a residual voltage of 0.5 to 2 V and a current of a few tenths of an ampere to several hundred amperes, solely limited by the maximum rated power dissipation. Power silicon thyristors (or silicon controlled rectifiers, SCRs) have operating currents of 1-2 kA and are used as contactors or current switches, in inverters, and rectified power supplies.

A thyristor turns on when the collector voltage is V_{ON} (the point of inflection between portions *1* and *2*). At $I_B \leqslant 0$, this voltage ranges from 25-50 V to 2-3 kV depending on the type of device used. As the positive base current (the triggering current) is increased, this voltage may fall to zero very nearly.

When a thyristor is turned off, the collector current falls to I_{OFF} ranging (at $I_B \geqslant 0$) from tens of microamperes to tens of milliamperes, depending on the type of device [3]. At $I_B \geqslant 0$, as usually happens in practice, for a thyristor to turn off it will suffice to reduce the operating current to $i_C \leqslant I_{OFF}$ by bringing down the collector voltage or raising the load resistance. Another way to turn off the device, as follows from the characteristics in Fig. 12.7*c*, is to reverse the polarity of the base current, $I_B < 0$. In existing thyristors, however, the "extinguishing" current is comparable with I_{OFF}; the ratio $I_{OFF}/I_{B,ext}$ being anywhere from 10 : 1 to 5 : 1 or even less [14].

The basic parameters of thyristors are the turn-on time, τ_{ON}, and the turn-off time, τ_{OFF}. The turn-on time is the sum of the delay time required to build up the excess minority-carrier charge in the collector junction and the time required for the barrier capacitance, C_b, to discharge. The turn-off time is the sum of the storage time required for the removal of minority-carrier charge and the time required for C_b to charge. The total turn-off time is usually by an order of magnitude longer than the turn-on time. For present-day low-power high-speed thyristors, τ_{ON} is tens of nanoseconds, and τ_{OFF} is hundreds of nanoseconds [14].

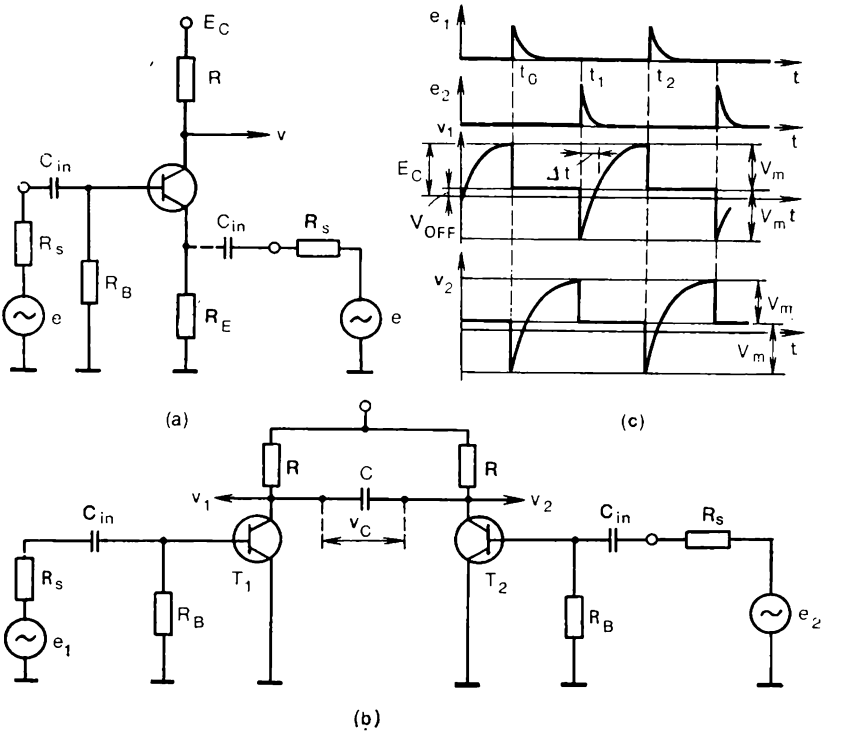


Fig. 12.8

Thyristor flip-flops. The circuit of a simple *RS* flip-flop built around a single trinistor is shown in Fig. 12.8*a*. The values of E_C and R are chosen such that

$$E_C < V_{ON},\ (E_C - V_{OFF})/R > I_{OFF} \tag{12.9}$$

Equation (12.9) locates three points of intersection between the load line and three regions of the thyristor's characteristic, that is, defines the conditions for the triggered operation of the circuit. If these conditions are satisfied, injection of positive and negative trigger pulses with an amplitude matching the characteristics of the device will cause the flip-flop to change state in the forward and backward direction, with a change of voltage equal to $V_m \approx E_C - V_{OFF}$. As already noted, such a flip-flop is substantially less sensitive to negative turn-off pulses than it is to positive pulses. Therefore, the latter may be fed into the emitter circuit, as shown by a dashed line in Fig. 12.8*a*.

A two-stage *RS* flip-flop (Fig. 12.8*b*) is more sensitive to trigger pulses and, also, is capable of delivering a noninverted and an invert-

ed output signal. To cause a change of state in the backward direction, each stage is coupled to the next by a capacitor C placed between the collector load resistors R. Suppose that at $t < t_0$ (Fig. 12.8*c*), $T1$ is OFF and $T2$ is ON (the first stable state of the flip-flop). Now the capacitor C is charged to $v_C = E_C - V_{resid}$, where V_{resid} is the residual voltage across $T2$, close to V_{OFF}. When a trigger pulse is applied at t_0, transistor $T1$ turns ON, and the negative change of collector voltage, $-V_m$, determined by the initial recharging current of capacitor C, is transferred to the collector of $T2$ which turns OFF in a time τ_{OFF} largerly controlled, as already noted, by the removal of excess minority-carrier charge from the collector junction. After capacitor C has recharged to $v_C = -E_C + V_{resid}$, the flip-flop finds itself in its second stable state ($T1$ is ON and $T2$ is OFF). The trigger pulse e_2 arriving at time t_1 initiates a similar chain of events which cause the flip-flop to go back to its first stable state. The minimum value of the coupling capacitor, C_{min}, should be chosen such that the time interval Δt (Fig. 12.8*c*) during which Eq. a negative potential is maintained at the collector of the OFF transistor is longer than its turn-off time τ_{OFF}:

$$\tau_{OFF} \leqslant \Delta t \approx 0.7CR \tag{12.10a}$$

In practice, it is always that $V_0 \ll E$. From Eq. (12.10*a*), we get

$$C_{min} \approx 1.4\tau_{OFF}/R \tag{12.10b}$$

Equations (12.9) and (12.10) define the conditions for the operability of the flip-flop. The resolving time, T_{min}, of such a flip-flop depends on the recharging time of the capacitor:

$$T_{min} = 3CR \text{ to } 4CR \tag{12.11a}$$

For $C = C_{min}$,

$$T_{min} \approx 4\tau_{OFF} \text{ to } 6\ \tau_{OFF} \tag{12.11b}$$

As a rule, $F_{max} = 1/T_{min}$ is tens or hundreds of kilohertz [30].

The circuit of Fig. 12.8*b* can be re-arranged into a T flip-flop by adding any one of the RC-networks for steered symmetrical triggering examined in Chap. 6.

REVIEW QUESTIONS AND PROBLEMS

12.1. Draw the basic circuit for a tunnel diode, using its volt-ampere characteristic; list the stable states and illustrate their use in pulse circuits.

12.2. Draw the basic circuit of a tunnel-diode multivibrator; plot its signal waveform; formulate the conditions for correct operation; determine the period of oscillation; explain what triggered operation is like.

12.3. Draw the circuit of a single-stage tunnel-diode flip-flop. Explain operation of the flip-flop in two modes.

12.4. Draw the circuit of an avalanche-transistor multivibrator; explain its operation; plot the signal waveform, and determine the period of oscillation for free-running operation.

12.5. Draw the circuit of a single-stage trinistor flip-flop; explain its operation.

12.6. Draw the circuit of a two-stage trinistor RS flip-flop; explain its operation; plot its signal waveform, and determine the resolving time.

13 FUNCTIONAL PULSE AND DIGITAL CIRCUITS

13.1. GENERAL

At present, the design of complex electronic circuits (systems) widely utilizes the "building-block" principle. It is based on the fact that information processing can be broken down into a sequence of a relatively small number of standard operations (microoperations). The execution of each microoperation can be assigned to a particular functional or operational circuit or unit. Some examples of standard microoperations are data selection, reception, storage and retrieval, comparison of values, arithmetical operations, pulse counting, and data conversion for subsequent display.

The various functional circuits or units are usually arranged in an hierarchical system in which the circuits at a lower level (the simpler units) may be part of a higher-level circuit or unit.

It is usual to call "digital" a functional circuit or unit which utilizes binary variables for input, intermediate results, and output.

A functional circuit or unit is called "pulse", if, in addition to binary variables, it utilizes, say, the amplitude of a pulse or the instantaneous value of a varying voltage to convey information.

13.2. REGISTERS

As already noted, an elementary unit of information in digital systems is a binary variable, or a *bit*. To handle numbers, characters, etc., one has to use bits arranged into what are called words—ordered strings of bits. In a word, the bits take up places or positions which are numbered by integers from 1 to n, where n is equal to the total number of bits in a word and defines the *word length*.

It is customary to use the capital letters A, B, X, etc. to designate words, and the low-case letters a_i, b_i, x_i, etc. to designate the respective bits (or the stages of the circuit), where the i's are the ordinal numbers of the places (or stages).

Words are held (stored), while information is being processed, in registers. Each bit of a word is written into a separate register stage which is basically a flip-flop. The stages of a register are numbered in the same order as the bits in a word.

From registers, words are usually transferred to combinational logic circuits where logical operations are performed on them. The

outcomes of these operations, in the form of one or several words, are written into output registers. If a register is built from internal-delay flip-flops, the result of an operation can be transferred back into the registers storing the operands (the words being operated upon).

Registers may be *parallel*, *serial*, or *serial-parallel*, according as words are written in. In a parallel register, all the bits of a word are written into all the register stages concurrently. In a serial register, each word is written one bit at a time, starting with that in the most or least significant position. As their name implies, serial-parallel registers are adapted to both modes of operation. They serve as a basis for the conversion of serial to parallel information and back.

Parallel Registers

A parallel register is an ordered set of flip-flops having common control and sync signals and individual data inputs to accept all the bits of the word being written in at the same time. In the general case, each stage of a parallel register has as many data inputs as there are information sources connected to it.

Data read-in and read-out in parallel registers usually occur at different times. Therefore, the register stages may be simple flip-flops.

Depending on the number of lines feeding data into the register stages, there are true-input and true-and-complement-input parallel registers. In a true-input parallel register, each bit of an input word is written in over a single-wire circuit which carries the signal representing the true value of the bit. A true-and-complement-input parallel register has a two-wire circuit to carry two signals, one representing the true value of each bit and the other, its complement.

A further subdivison of parallel registers is into *asynchronous* and *synchronous*. In a synchronous parallel register, control signals are gated ("sampled" or "strobed") by clock pulses. In this way, the register is rendered less immune to false pulses accompanying the valid control signals. It should be noted, though, that within an asynchronous parallel register the control signals act in effect as clock pulses, as they control the order in which the flip-flops are switched in its bit stages.

Consider several typical circuits of parallel register stages.

Clear/enable true-input parallel register. The register has data inputs $A_1, A_2, \ldots, A_m$ and control inputs $R', V_1, V_2, \ldots, V_M$. In writing a word A_k into the register, one applies a reset (clear) pulse to input R', and an enable pulse to enable input V_k. The first pulse clears (resets to zero) the register, and the second causes the word A_k to be written into the register.

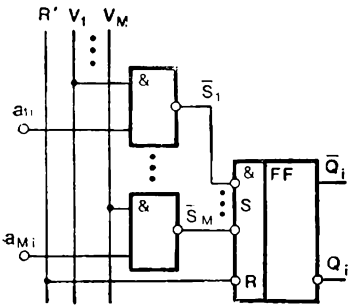


Fig. 13.1

Figure 13.1 shows a logic diagram of a NAND gate register stage. The stage uses a flip-flop which has an increased number of complementing set inputs, S, and there is a proportionally increased number of input NAND gates.

In a synchronous parallel register, the enable signals V_1, V_2, . . . , V_k are gated (sampled) by a clock pulse applied simultaneously to all the input NAND gates (see Fig. 13.1).

Enable true-input parallel register. The register has a number M of data inputs A_1, A_2, . . ., A_M and a number M of enable inputs V_1, V_2, . . ., V_M. A word A_k is written into the register in response to a signal applied to input V_k.

The circuit of such a register stage based on D-type latch flip-flops is shown in Fig. 13.2*a*. To have a larger number of D inputs a distributed (or wired) AND connection is used. This is an economical configuration because its enable-pulse inverters are common to the entire register. The enable signal is a binary 0 applied to input

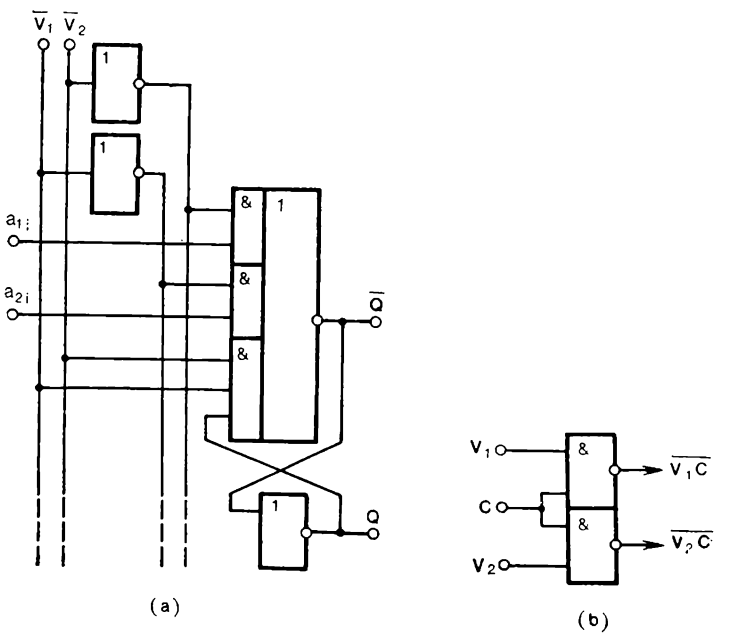


Fig. 13.2

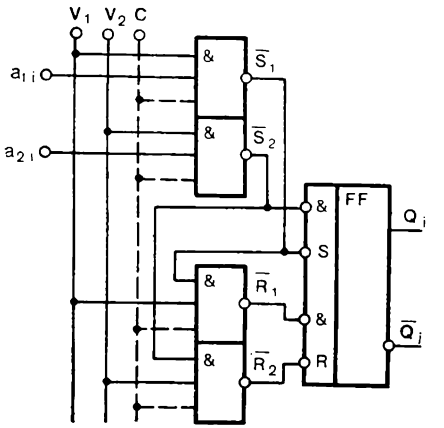


Fig. 13.3

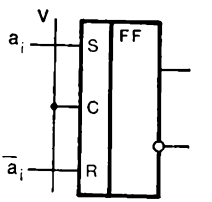


Fig. 13.4

$\overline{V}_1$ or $\overline{V}_2$. Enable signals are gated by clock pulses, using input NAND gates (Fig. 13.2*b*).

Figure 13.3 shows the circuit of a NAND gate register stage. The register is similar in structure to a *D* flip-flop with set-reset inputs. In order to obtain a larger number of *D* inputs tied together, an increased number of $\overline{R}$ and $\overline{S}$ inputs is provided on the flip-flop circuit along with an increased number of pairs of input NAND gates.

The manner in which the stage of a *D* flip-flop register is clocked is shown by dashed lines in Fig. 13.3.

True/complement-input register. Each stage of such a register is a clocked *RS* flip-flop which can be made with NAND or AND-NOR gates. For simplicity, Fig. 13.4 shows only one register stage. A larger number of inputs can be had by providing an increased number of gate pairs at the flip-flop input (in about the same way as this is done in the *D* flip-flop examined earlier). Clock signals (not shown in the diagram) can be applied by using a clock input *C* made common to all the AND gates (the distributed or wired AND connection).

Serial Registers

Serial registers provide facilities to shift a word one place to the right or left, or both. Quite aptly, they are called *shift* (or *shifting*) *registers.* The microoperation of a right shift moves every bit so that it takes up the next higher significant position; taking the *k*th bit as an example, a right shift will move it to the $(k + 1)$st position. Similarly, a left shift moves every bit so that it takes up the

next lower significant position; taking the kth bit as an example again, a left shift will leave it in the $(k-1)$st position.

As already noted, a shift register can move the bit string stored to only the left or right, or both. Consider a shift register as a sequential circuit and define its transition functions for the various form of shifting.

The internal state of the register is specified by the sum of the states of its stages, $Q_1, Q_2, \ldots, Q_n$, which give in fact the value of the word stored in the register.

Unidirectional shift register. It has two data (D) inputs, two control (V) inputs, and one output, Y. The transition and output functions of the register as a sequential circuit are defined by switching functions as follows:

—for a right-shift register

$$\begin{aligned} Q_k(t+1) &= VQ_{k-1}(t) + \bar{V}Q_k(t) \quad k = 1, 2, \ldots, n \\ Q_0 &= D \\ Y &= Q_n \end{aligned} \tag{13.1}$$

—for a left-shift register

$$\begin{aligned} Q_k(t+1) &= VQ_{k+1}(t) + \bar{V}Q_k(t) \quad k = 1, 2, \ldots, n \\ Q_{k+1} &= D \\ Y &= Q_1 \end{aligned} \tag{13.2}$$

It follows from the above relations that a shift is performed when $V = 1$. The bit position vacated by the shift is taken up by an input variable D, and the bit shifted out of the register appears at its output which is the outermost cell towards which the shifting is done. When $V = 0$, the word stored in the register remains unchanged.

Bidirectional shift register. This type of register has two data inputs, D_R and D_L, and two data outputs, V_R and V_L. A bidirectional shift register may have two symmetrical control inputs V_L and V_R or one control input V and one sign input Sgn, depending on the form of control signals used.

The former type of bidirectional register operates as a right-shift register when $V_L = 0$ and as a left-shift register when $V_R = 0$. Application of both V_R and V_L control signals at the same time is forbidden. That is,

$$V_R V_L = 0 \tag{13.3a}$$

The transition function of this register has the form

$$\begin{aligned} Q_k(t+1) = V_k Q_{k-1}(t) + V_L Q_{k+1}(t) + \bar{V}_L \bar{V}_R Q_k(t) \\ k = 1, 2, \ldots, n \end{aligned} \tag{13.3b}$$

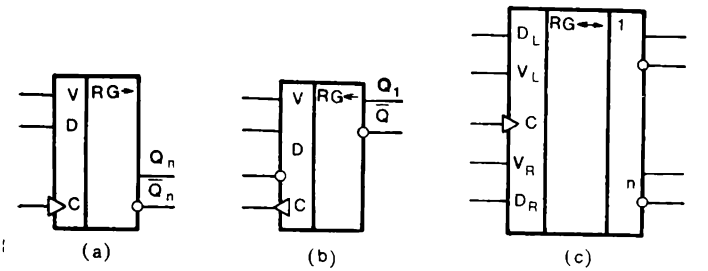


Fig. 13.5

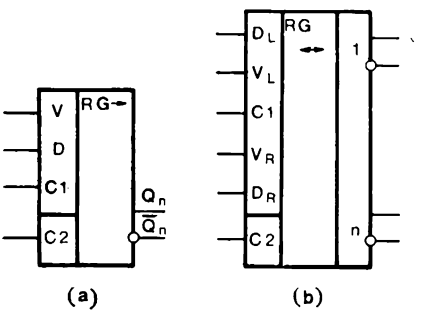


Fig. 13.6

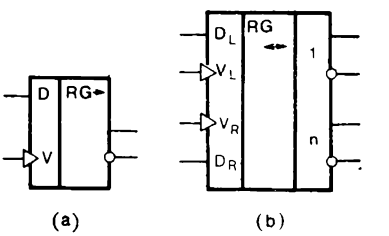


Fig. 13.7

The latter type of bidirectional register operates as a right-shift register when Sgn = 0 and as a left-shift register when Sgn = 1. Its transition function can be written

$$Q_k(t+1) = V(\overline{\text{Sgn}}\,\overline{Q_{k-1}}(t) \dotplus \text{Sgn}\,Q_{k+1}(t)) \dotplus \overline{V}Q_k(t) \quad k = 1, 2, \ldots, n \tag{13.4}$$

The inputs and outputs of bidirectional registers are written

$$Q_0 = D_R, \quad V_R = Q_n, \quad Q_{n+1} = D_L, \quad V_L = Q_1$$

A serial register may be built synchronous or asynchronous. For a synchronous serial register, the time unit, t, is the spacing between two successive clock pulses. A synchronous serial register may have one, two, three and more clock inputs. Most often, use is made of synchronous serial registers with one clock input (Fig. 13.5) and two clock inputs (Fig. 13.6).

An asynchronous serial register (Fig. 13.7) is a sequential circuit in which each operation is performed at discrete instants of time, following a change in the signal at the control (V) input. In Eqs.

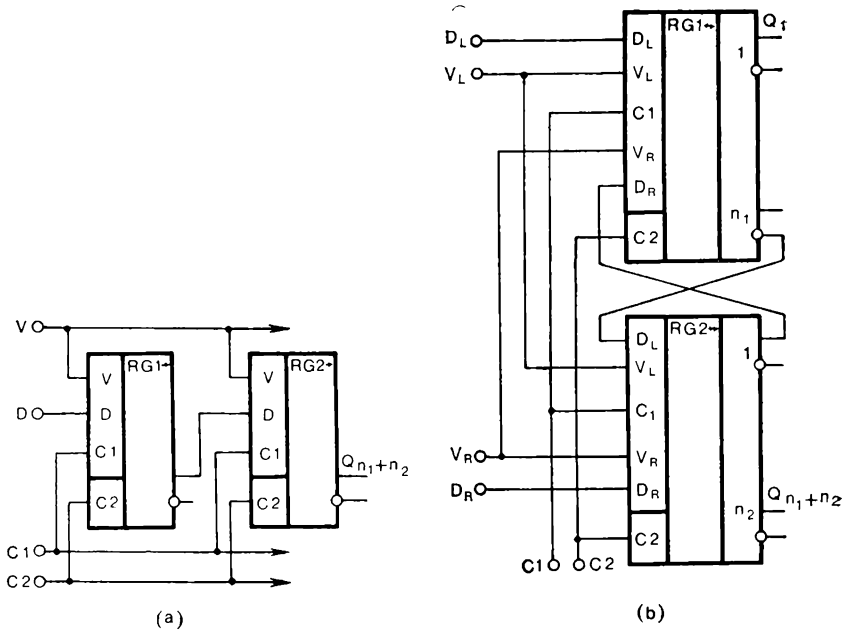


Fig. 13.8

(13.1) through (13.4), the control variables V, V_R and V_L take on a value of unity when the respective control input accepts the positive- or negative-going edge of the signal.

Serial registers in cascade. Registers are connected in cascade to increase their storage capacity. Figure 13.8*a* shows two unidirectional shift registers in cascade. If register $RG1$ can store n_1 bits, and register $RG2$, n_2 bits, the cascade will be able to store $n_1 + n_2$ bits of data. A cascade connection of bidirectional registers is shown in Fig. 13.8*b*.

A serial register with an arbitrary number of stages may be treated as a cascade of one-bit registers.

One-bit register. This is a sequential circuit with two internal states, $Q = 1$ and $Q = 0$. Its transition functions can be derived from Eqs. (13.1) through (13.4) on setting $k = 1$.

For a right- or left-shift register,

$$Q(t+1) = V(t)\,D(t) + \bar{V}(t)\,Q(t) \qquad (13.5a)$$

For a bidirectional shift register with two control inputs,

$$Q(t+1) = V_R D_R + V_L D_L + \bar{V}_L \bar{V}_R Q(t) \qquad (13.5b)$$

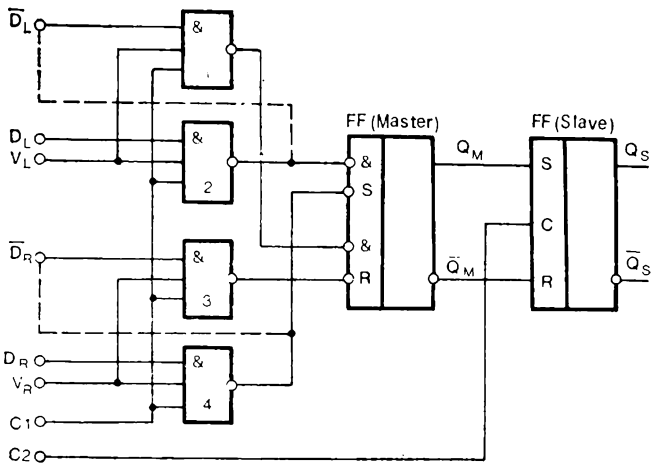


Fig. 13.9

For a bidirectional shift register with a sign input,

$$Q(t+1) = (D_R\,\overline{\text{Sgn}} + D_L\,\text{Sgn})\,V + \overline{V}Q(t) \qquad (13.5c)$$

From Eq. (13.5a) it is seen that a D flip-flop may well be used as a stage in a unidirectional register. Also, clocked RS and JK flip-flops operated as a true/complement D flip-flop can be used, where the S or J input accepts the true D value, whereas the R or K input, its complement.

In the former case, the stage of a bidirectional shift register is a D flip-flop with two groups of inputs. Figure 13.9 shows the stage of a clear/enable bidirectional shift register. The circuit is in effect a master-slave flip-flop (see Chap. 6). The master flip-flop is a primitive D flip-flop with a single clock input and true/complement data inputs, for which the set-reset functions may be written

$$s = D_L V_L C_1 + D_R V_R C_1$$
$$r = \overline{D}_L V_L C_1 + \overline{D}_R V_R C_1$$

As is seen, at clock time $C1$, the master flip-flop goes to the state $Q_M = D_L$ if $V_L = 1$, and to the state $Q_M = D_R$ if $V_R = 1$. When $V_L V_R = 0$, the set and reset functions are both zero, and the state of the flip-flop remains unchanged.

The slave flip-flop is a simple RS flip-flop with a single clock input. When the clock signal equals 1, the slave flip-flop sets to the state $Q_S = Q_M$.

Using the connections shown by dashed lines, one can build a register stage from single-phase D flip-flops.

In order to obtain a similar stage for a unidirectional shift register,

it will suffice to remove one pair of input NAND gates (*1* and *2* or *3* and *4*) in the circuit of Fig. 13.9.

As regards circuit configuration, enable-only synchronous serial registers are close to asynchronous serial registers. As will be shown shortly, an asynchronous circuit can be converted into a synchronous circuit by replacing the control pulse V at the input of the asynchronous register with a logic product VC. In other words, the task of the clock signal is to strobe ("sample") the control signal V.

The transition function for an asynchronous one-bit register can be derived from Eq. (13.5), with $V\mathrm{d}V$ replacing V.

For a unidirectional shift register,

$$Q(t+1) = DV\,\mathrm{d}V + Q\overline{V\,\mathrm{d}V}, \qquad \mathrm{d}Q = (D \oplus Q)\,V\,\mathrm{d}V \tag{13.6a}$$

For a bidirectional shift register with two symmetrical control inputs,

$$\begin{gathered} Q(t+1) = D_L V_L\,\mathrm{d}V_L + D_R V_R\,\mathrm{d}V_R + Q\overline{V_R\,\mathrm{d}V_R V_L\,\mathrm{d}V_L} \\ \mathrm{d}Q = (Q \oplus D_L)\,V_L\,\mathrm{d}V_L + (Q \oplus D_R)\,V_R\,\mathrm{d}V_R \end{gathered} \tag{13.6b}$$

For a bidirectional shift register with a control input and a sign input,

$$\begin{gathered} Q(t+1) = (D_L\,\mathrm{Sgn} + D_R\overline{\mathrm{Sgn}})\,V\,\mathrm{d}V = Q\overline{V\,\mathrm{d}V} \\ \mathrm{d}Q = [(D_L\,\mathrm{Sgn} + D_R\overline{\mathrm{Sgn}}) \oplus Q]\,V\,\mathrm{d}V \end{gathered} \tag{13.6c}$$

The above relations apply when the stage changes state in response to a negative control signal V; so the V input is a complementing dynamic input. It is possible to configure a stage with a true dynamic V input. Then the variable V in Eqs. (13.6) must be replaced by $\overline{V}$.

A likely arrangement of a stage for an asynchronous unidirectional shift register and its logic symbol are shown in Fig. 13.10. It is essentially a D flip-flop with a true/complement data (D) input. Initially, $V = 1$, the input gates *1* and *2* are enabled, and the RS flip-flop *5* is set to a state $Q' = D$. The value of Q' is transferred

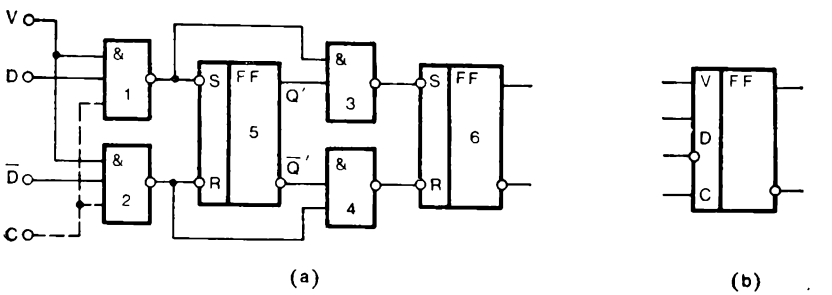


Fig. 13.10

to the output RS flip-flop, 6, as V goes from 1 to 0. Thus, the V input in this arrangement is a complementing dynamic input.

A multibit register formed by a cascade of the above type of asynchronous devices will operate properly if the V signal goes to all of its stages at the same time. This might be done by taking the V signal from a common logic gate. However, the fan-out of existing logic gates is relatively low, and the resultant multibit register could only have a limited number of stages. It would seem that the situation could be rectified by including a power stage to drive the inputs of all the stages in a multibit register. Unfortunately, the signal-steering circuits would then carry substantial currents which might give rise to strong inductive pickup.

In practice, control signals are steered to the register stages by means of a multitier tree structure of logic gates (usually inverters) (Fig. 13.11). If the fan-out of one logic gate is N, then the number of gates in the next tier will be N times that in the previous tier. So, an n-tier circuit can be loaded into N^n logic gates or $0.5\,N^n$ register stages. The instants when the control signals V' are changed at the input of the multitier tree structure are delayed from the initial change in the variable V by the time required for the signal to propagate down the chain of logic gates. Even in circuits built of identical logic gates, the discrepancy in switching time may be considerable (by a factor of 4 to 7 or more). Because of this, the control signals do not act on their respective register stages simultaneously. For the circuit configuration in Fig. 13.10, this might lead to ambiguous performance of the register usually known as the race condition.

For better insight into the race problem, consider the operation of two adjacent register stages, the ith and the $(i + 1)$st, assuming that the control signal applied to the input of the $(i + 1)$st stage changes later than the control signal applied to the ith stage (Fig. 13.12). Suppose also that initially the ith stage awaits a signal

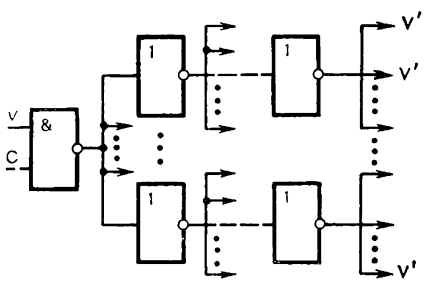


Fig. 13.11

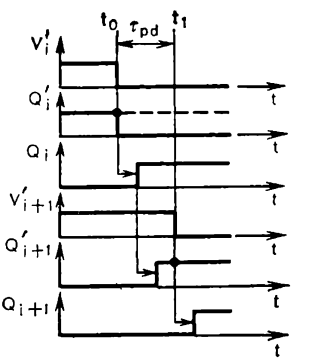


Fig. 13.12

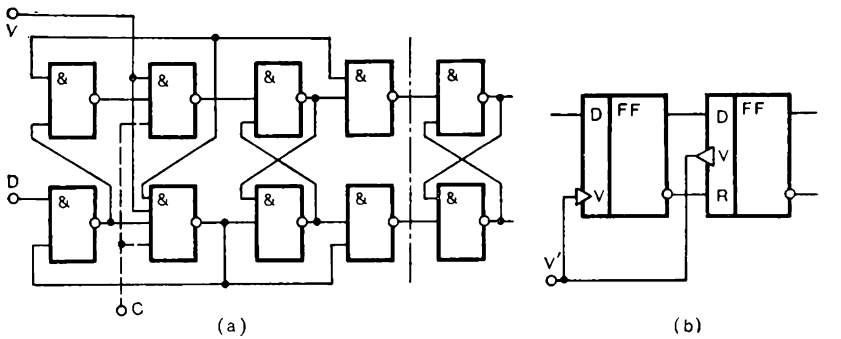


Fig. 13.13

that will drive it to the 1 state, $Q_i' = 1$, and the $(i + 1)$st stage awaits a signal that will drive it to the 0 state, $Q_{i+1}' = Q_i = 0$. At time t_0, the negative-going edge of V_i causes the ith stage to change state. Because the negative-going edge of the V_{i+1}' signal is delayed from t_0, the $(i + 1)$st stage will remain in the previous state. During this time delay, τ_d, Q_{i+1}' might change again, so that in response to the V_{i+1} signal, the $(i + 1)$st stage might go to a 1, rather than a 0 state, which is a malfunction.

There are different ways to avoid the race problem. One is to employ a lockout circuit (see Sec. 6.6) that will forbid another change of state by the Q_{i+1}' stage during the delay. Figure 13.13*a* shows the circuit of a register stage which, like that in Fig. 13.10, has two memory cells and a circuit that forbids data entry into the second cell. Operation of the stage can be modelled by a cascade connection of two flip-flops one of which has a true dynamic V input, and the other, a complementing dynamic input (see Fig. 13.13*a*).

A disadvantage of the above configuration is that it uses a larger count of elements than the circuit in Fig. 13.10. This can be avoided by combining functionally the output flip-flop with the input gates of the next stage. This principle is embodied in the multibit register shown in Fig. 13.14. The first stage has a true input to simplify its interfacing with other digital circuits. The remaining stages use true/complement inputs. Each stage contains four NAND logic gates.

Consider the transfer of a signal from the kth stage to the $(k + 1)$st stage of this register. Assume that initially the control signals are $V_k' = V_{k+1}' = 0$. Gates *3* and *4* are disabled, and the state of the stages is determined by that of flip-flops *5* and *6*. To make our discussion more specific, let the kth stage be in the 1 state, $Q_k = 1$, and the $(k - 1)$st stage in the 0 state, $Q_{k-1} = 0$. Because gates *7* and *8* are enabled, input flip-flops *1* and *2* of the $(k + 1)$st stage are

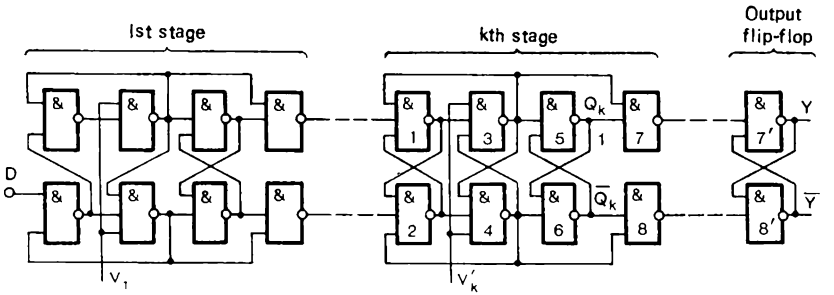


Fig. 13.14

in the 1 state, $Q''_{k+1} = Q_k = 1$, and, similarly, $Q''_k = Q_{k-1} = 0$. Consider the transients occurring as V goes from 0 to 1. Two cases may arise.

1. The positive edge of the V'_k signal leads that of the V'_{k+1} signal. Because gates *7* and *8* are disabled, the change of state in flip-flops *5* and *6* cannot be transferred to the input flip-flops, *1* and *2*, of the next stage. Therefore, the input cell of the $(k + 1)$st stage remains in the 1 state, $Q''_{k+1} = 1$. In fact, this cell serves to "memorize" the state of the kth stage for the duration of the time delay in the positive edge of the V'_{k+1} signal. With arrival of the positive edge, the 1 state of the input flip-flops is transferred to flip-flops *5* and *6* and, at the same time, produces a signal that disables gates *1* and *4*.

2. The positive edge of the V'_k signal lags behind that of the V'_{k+1} signal. Now, the 1 state of flip-flops *1* and *2* in the $(k + 1)$st stage is transferred to flip-flops *5* and *6* before gates *7* and *8* in the kth stage are disabled. The lockout signal generated in the $(k + 1)$st stage confirms the 0 signal from the output of gate *7*. After V'_k goes to 1, both outputs of the kth stage are set to 1, and the 0 state of flip-flops *1* and *2* is written into flip-flops *5* and *6*.

Now we shall examine the events that take place when the control signal changes in the reverse direction. As before, two cases are likely to happen.

1. The negative edge of the V'_k signal leads that of the V'_{k+1} signal. As gates *7* and *8* are enabled, this causes a 0 signal to be applied to the input gate *2* of the $(k + 1)$st stage, thereby switching it to the 1 state. The signal will not propagate any farther in the $(k + 1)$st stage owing to the lockout action of gates *1* and *4*. After the V'_{k+1} signal changes from 1 to 0, gates *3* and *4* are disabled, gate *1* is enabled and flip-flops *1* and *2* are set to a 0 state, $Q''_{k+1} = Q_k = 0$.

2. The negative edge of the V'_{k+1} signal lags behind that of the V'_k signal. As gates *3* and *4* in the $(k + 1)$st stage are disabled, and flip-flops *5* and *6* are set to the 1 state, $Q_{k+1} = 1$.

In this way, the state of the kth stage is correctly transferred to the $(k + 1)$st stage despite any discrepancy in switching time between the V'_k and V'_{k+1} signals.

A distinction of the above arrangement is that at $V = 1$, an indeterminate 1 state appears at both the true and complementing outputs. To remove this indeterminacy in the output, recourse is had to an additional RS flip-flop with complementing inputs, made with two NAND gates.

A bidirectional shift register with a sign input can be built in a similar manner. In this case, the stage must include an additional pair of gates $7'$ and $8'$ to interface the output of the $(k + 1)$st stage with the input of the kth stage. When Sgn $= 0$, gates 7 and 8 are disabled, whereas at Sgn $= 1$, gates $7'$ and $8'$ are disabled.

Let us dwell in brief on the structure of a synchronous enable-only shift register. For simplicity, we shall limit ourselves to a unidirectional shift. Assuming that the discrete (sampling) times occur at the negative edge of the clock signal, the transition function of the asynchronous D flip-flop modelling the operation of a register stage may be written

$$\mathrm{d}Q = (D \oplus Q)\, VC\, \mathrm{d}C \tag{13.7}$$

In practice, registers are configured, assuming that the control signal V changes between two successive clock pulses, that is, $C\mathrm{d}V = 0$ and $\mathrm{d}C\mathrm{d}V = 0$. Then the following equalities hold

$$\mathrm{d}\,(CV) = (C \oplus \mathrm{d}C)\,(V \oplus \mathrm{d}V) \oplus \bar{C}V = V\,\mathrm{d}C$$
$$(CV)\,\mathrm{d}\,(CV) = CV \cdot V\,\mathrm{d}C = VC\,\mathrm{d}C$$

Hence, the transition function, Eq. (13.7), may be replaced by an equivalence function

$$\mathrm{d}Q = (D \oplus Q)\, CV\, \mathrm{d}\,(CV) \tag{13.8}$$

As follows from Eq. 13.8, a synchronous enable-only register can be implemented as an asynchronous register controlled by the logic product of the clock pulse C and the control signal V. In this arrangement, the control signal is gated ("sampled") by a synchronizing pulse. Several ways of injecting the clock signal are shown in Figs. 13.10, 13.13, and 13.14 by dashed lines.

Parallel-Serial Registers

These registers combine the properties of parallel and shift registers. The core of a parallel-serial register is a shift register which is extended to include parallel read-in circuits. As an example, we shall examine the circuit in Fig. 13.15. It is built around a NAND

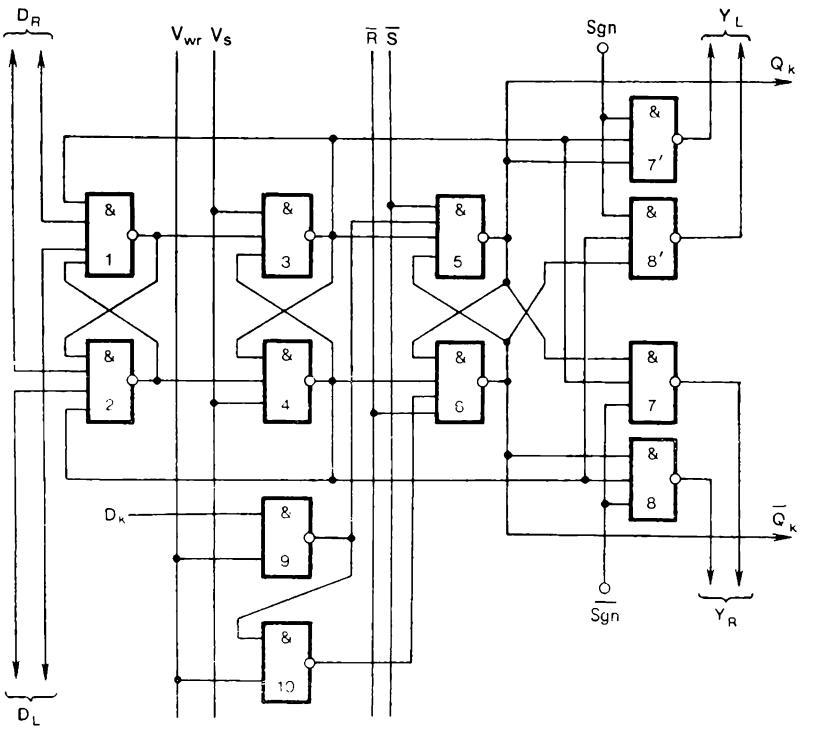


Fig. 13.15

gate asynchronous bidirectional sign-input shift register. For parallel read-in, the circuit is supplemented with input gates *9* and *10* which form, between them, a steered single-phase D input. Also, it has set-reset inputs, $\overline{S}$ and $\overline{R}$, which are used to set all register stages to a 0 or 1 state initially. Shifting is done in response to a pulse applied to the V_s input, and parrallel writing in response to a pulse applied to the V_{wr} input.

Dynamic Registers

The registers we have examined so far are static registers. That is, the bits of the word stored remain in the same register stages throughout the storage time. Apart from static registers, digital circuits and systems widely employ dynamic registers.

The dynamic register in Fig. 13.16 is a loop or ring structure in which the stored word is continuously circulating. Quite aptly, this form of register is often called a recirculating register. It consists essentially of a unidirectional shift register and an input logic

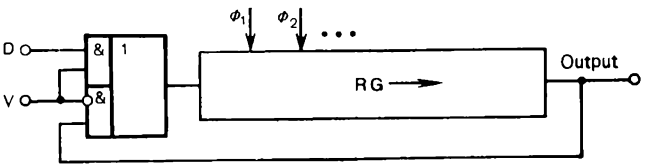


Fig. 13.16

circuit which serves to select the desired mode of operation—read-in of new information or regeneration of the information coming from the output—in response to an enable signal V. The shift register proper has no enable input, and the shift operation is performed continuously at the repetition frequency of clock pulses. If the clock rate is sufficiently high, the storage medium in dynamic shift registers may be LC networks. More frequently, however, they are based on MOS-transistors in which the storage capability is provided by internal capacitances (the gate-to-source capacitance C_{GS} and the parasitic capacitance between the electrodes and substrate). Among the advantages of dynamic MOS shift registers are simplicity, high speed, low power dissipation, and high manufacturability.

Dynamic MOS shift registers usually use a multiphase clock, two- and four-phase types being most frequent.

A typical stage of a two-phase MOS dynamic shift register can be obtained by cascading two MOS inverters (Fig. 13.17*a*). Here, $T1$ and $T4$ are inverters proper, $T2$ and $T5$ constitute the load, whereas $T3$ and $T6$ act as bidirectional switches. Clock waveforms Φ_1 and Φ_2 (Fig. 13.17*b*) are applied to the gates of the load and switch MOSFETs.

When $\Phi_1 = \Phi_2 = 0$, $T2$, $T3$, $T5$, and $T6$ are OFF, and the register draws no power from the supply. The parasitic capacitances C_1 and C_2 slowly discharge due to the action of reverse currents across the source-substrate pn-junctions of the OFF switch transistors. The maximum spacing between clock pulses is chosen such that

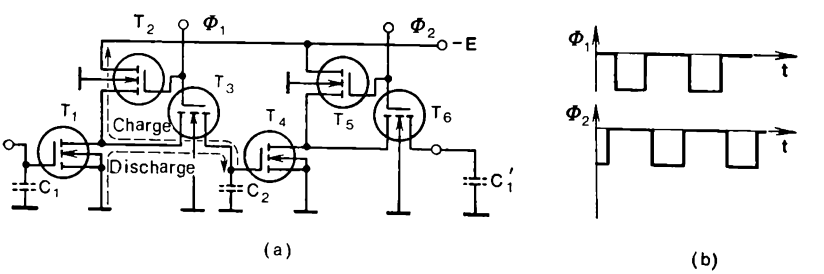


Fig. 13.17

a drop in the voltage across the parasitic capacitances could not cause the inverters to change state.

Applying a clock pulse Φ_1 turns on $T2$ and $T3$. If $T1$ is held OFF by a high level on C_1, C_2 will charge via $T2$ and $T3$ towards a low (negative) voltage level. If, on the other hand, $T1$ is ON, it will offer a low resistance, C_2 will discharge via $T1$ and $T3$ and pull up to a high (nearly zero) level. In this way the bit of data is transferred from C_1 to C_2 and the logic level deposited on C_2 in response to the first clock pulse Φ_1 is the complement of the logic level on C_1.

Applying a second clock pulse Φ_2 causes the bit of information to be farther transferred from C_2 to C_1' in a similar manner so that the level stored on C_1' is the same as existed on C_1 before the application of the clock waveforms, but delayed by a clock period. As a result, the string of bits stored is shifted one position.

A limitation of a two-phase MOS dynamic shift register is that when the output of any of the two inverters goes high, its load and inverting transistors are conducting together. This constitutes a heavy drain on the power supply. Also, as in static registers, the transconductance of the inverter transistor must exceed that of the load transistor. As a consequence, they will take up a larger area on a chip.

A better choice is a four-phase MOS dynamic shift register. One stage of such a register is shown in Fig. 13.18*a*. The bit of information stored on C_1 is transferred to C_2 in two steps (Fig. 13.18*b*). First, the clock pulse Φ_1 turns on $T3$, and C_2 charges up to the supply voltage, $-E$. Next, the clock pulse Φ_2 turns on $T2$ and, if $T1$ has been ON (C_1 has stored a low logic level), C_2 discharges to zero; otherwise, it remains at a low level. Transfer of a bit of information from C_1 to C_1' timed by clock pulses Φ_3 and Φ_4 proceeds in a similar manner.

Transistors $T2$ and $T3$ (or $T5$ and $T6$) are never conducting together. This results in a saving of power which is solely used up to

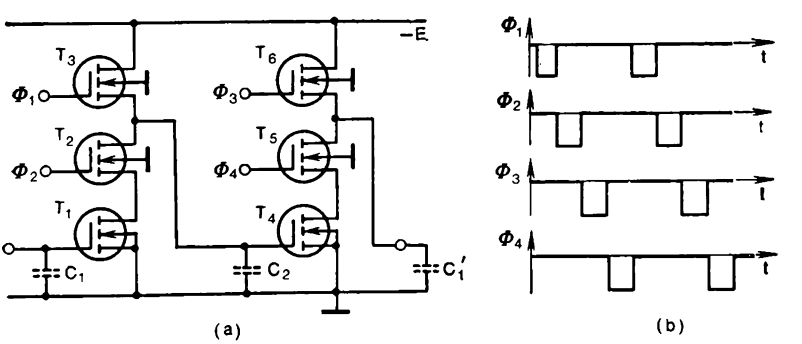


Fig. 13.18

charge the capacitances. Also, the transistors used in the circuit may be all of the same type and rating, which fact simplifies the circuit design and reduces the area occupied on a chip.

Charge-Transfer Devices (CTD) Shift Registers

An extension of the ideas underlying MOS dynamic shift registers has led to the invention of novel types of MOS ICs which have come to be known as charge-transfer devices (CTDs). We shall dwell on two forms of CTDs, namely bucket-brigade devices (BBDs) and charge coupled devices (CCDs).

BBDs (Fig. 13.19*a*). Physically this type of device is an IC structure (Fig. 13.19*b*) consisting of *p*-regions produced by diffusion in the surface layer of an *n*-type substrate, and a system of metal electrodes separated from the semiconductor material by an oxide layer. Two adjacent *p*-regions form a transistor, one of the *p*-regions being common to two series-connected transistors. In one, it acts as the drain, and in the other, as the source. At the same time, this *p*-region is one plate of a storage capacitor. The other capacitor plate is a metal electrode which also doubles as the transistor gate. As is seen, each element of the IC structure acts in several capacities. This leads to a high packing density which is a major advantage of these devices.

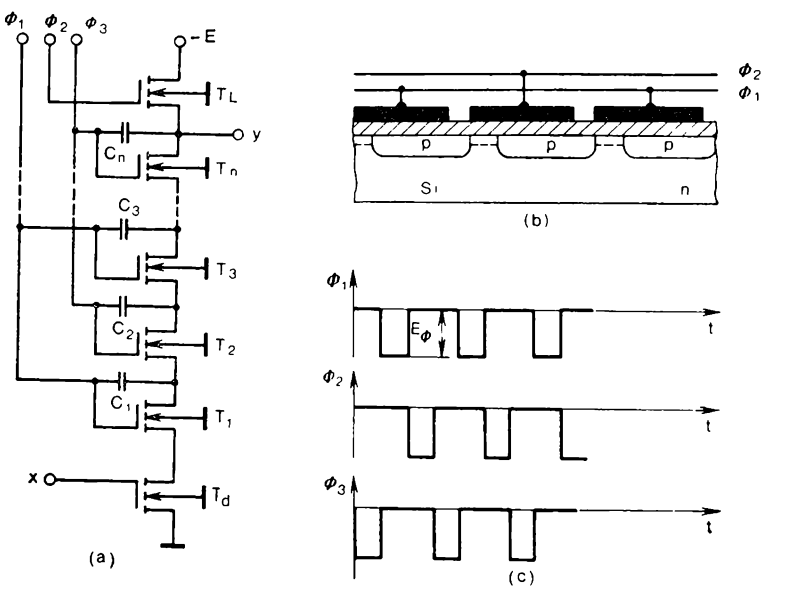


Fig. 13.19

The register is clocked by three-phase waveforms Φ_1, Φ_2, and Φ_3, shown in the diagram of Fig. 13.19*c*. Assuming that the driver input transistor T_d is OFF, application of the clock waveforms to the respective transistors (T_1, T_2, . . ., T_n, and T_L) will turn them on one after another, and the storage capacitors C_1 through C_n will charge up, whereas the source and drain voltages of the transistors will go down. The capacitors will cease charging when these voltages have become as low as $-(E_\Phi - V_{GS,\,th})$, where E_Φ is the amplitude of the clock waveform and $V_{GS,\,th}$ is the threshold voltage of the transistors. No further fall in the voltages is possible, because the applied clock pulses leave the respective transistors just at the verge of turn-on. The resultant voltage is in fact the low logic level

$$V^L = -(E_\Phi - V_{GS,\,th}) \tag{13.9}$$

Now suppose that, by the time when the clock pulse Φ_1 is applied, the driver transistor is turned on by an input signal. Then the current flowing in T_d and T_1 will charge C_1, thereby incrementing the drain voltage of T_1 by ΔV. This increment usually exceeds the difference $E_\Phi - V_{GS,\,th}$. So, when the clock pulse ceases, the drain-substrate *p-n* junction is rendered conducting, and C_1 drops to a nearly zero voltage representing the high logic level, $V^H \approx 0$. Now, we may take it that the total increment in the voltage on C_1 is equal to $E_\Phi = V_{GS,\,th}$.

The next clock pulse Φ_3 causes the charge on C_1 to be transferred to C_2, so that the drain voltage of T_1 goes back to $V^L = -(E_\Phi - V_{GS,\,th})$ and the drain voltage of T_2 is incremented by ΔV. In a similar manner, the charge is transferred from C_2 to C_3, etc. Each time a clock pulse Φ_1 is applied, C_1 goes to the low logic level, V^L, and, depending on the state of the driver transistor, C_1 may or may not receive an additional charge during the clock pulse Φ_1; this charge will then be moved along the chain of capacitors C_2 through C_n. The magnitude of the output signal appearing during Φ_1 is determined by the voltage on C_n. The output voltage changes in two steps. During the first step, Φ_2 drives C_n to a low logic level. During the second, Φ_3 causes the signal to be transferred from C_{n-1} to C_n. Taking Φ_1 as the time origin, the output signal may be written

$$y\,(t + n/2) = \overline{x\,(t)}$$

whence it follows that the network in question is equivalent to an $n/2$-bit shift register.

Let us trace in more detail the transfer of charge from C_{i-1} to C_i in response to $\Phi_{1(3)}$. An equivalent circuit for the charge transfer path is shown in Fig. 13.20. In the equivalent circuit transistor T_i through which the charge is transferred is represented by a current generator i_D. Initially (prior to application of a clock pulse), v_i

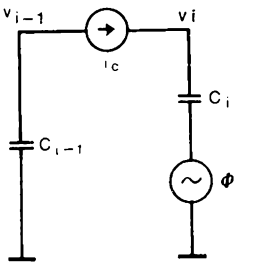


Fig. 13.20

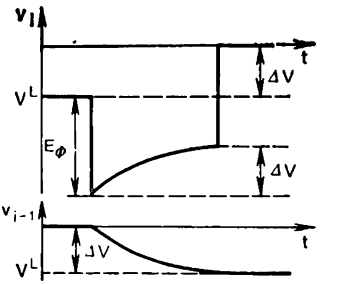


Fig. 13.21

is equal to the low logic level, V^L, and v_{i-1} differs from V^L by ΔV which determines the magnitude of the signal being transferred (Fig. 13.21). Applying a clock pulse turns on the transistor. At the same time, its drain voltage undergoes a step change which drives the transistor into the gently sloping region of its volt-ampere characteristics. The drain current is

$$i_D = \nu (v_{GS} - V_{GS,\,th})^2 = \nu (v_{i-1} - V^L)^2 \qquad (13.10)$$

This current substains the transfer of the positive charge from C_{i-1} to C_i until v_{i-1} falls to V^L when $i_D = 0$. If C_{i-1} and C_i are equal, v_{i-1} and v_i will be incremented by the same absolute amount. Therefore, at the end of the clock pulse, v_i becomes equal to $V^L + \Delta V$. In other words, ΔV is transferred from C_{i-1} to C_i.

The transient change in v_{i-1} can be written

$$C \frac{dv_{i-1}}{dt} = -i_C \qquad (13.11)$$

where C is the capacitance of the storage capacitor.

Solving Eqs. (13.10) and (13.11) simultaneously, subject to the initial condition $v_{i-1}(0) = V^H$, we obtain

$$v_{i-1}(t) = V^L + \left(\frac{1}{V^H - V^L} + \nu t/C\right)^{-1} \qquad (13.12)$$

As follows from Eq. (13.12), in the course of charge transfer v_{i-1} obeys a hyperbolic law and approaches V^L asymptotically. For v_{i-1} to be close to V^L at the instant when the pulse ceases ($t = \tau$), it is essential to satisfy the condition

$$\nu\tau/C \gg 1/(V^H - V^L)$$

under which $v_{i-1}(\tau)$ is approximately equal to $V^L + C/\nu\tau$. The term $C/\nu\tau$ represents the dynamic losses suffered by the signal as it is transferred through the cell structure.

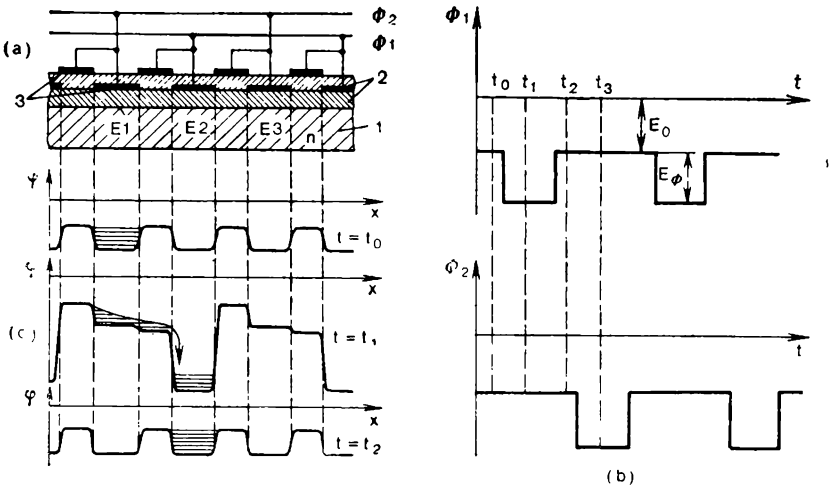


Fig. 13.22

Charge-coupled devices. Imagine a MOS transistor with an extremely long channel and with many gate electrodes closely spaced between source and drain. Each gate electrode and the substrate (*n*-type silicon) form a MOS capacitor which can store charge. This charge is accumulated when a sufficiently high negative voltage is applied to an electrode, so that a potential energy well (a depletion region) is induced under it. If now this voltage is removed from the electrode in question and simultaneously applied to that next down the channel, the charge packet will move in the same direction to fill the potential well induced next. By repeating this process, charge is transferred from well to well, and hence this configuration is called a *charge-coupled device**. Extremely high-density shift registers can be built with CCDs.

Figure 13.22*a* shows a fragment of a two-phase *p*-channel CCD** along with the clock waveforms Φ_1 and Φ_2 it uses in Fig. 13.22*b*. The potential profile (variations in potential φ with distance x) in the substrate at times t_0, t_1, and t_2, differing in the relative amplitudes of Φ_1 and Φ_2 is shown in Fig. 13.22*c*. The potential wells are formed under the gates lying closer to the substrate surface. Suppose that at time t_0 the potential well under electrode $E2$ is empty and that under $E1$ holds a charge packet of minority carriers (holes), Q (see Fig. 13.22*c*). Application of a negative clock pulse Φ_1 (at

* Some authors call it a charge-transfer device (CTD).—Translator's note.

** The name "p-channel" is used because the charge packet consists of holes in the *n*-type substrate.—Translator's note.

$t = t_1$) deforms the potential profile so that the charge "overflows" from the well under $E1$ into the well under $E2$. At the same time, the potential well is vacated under $E3$ in readiness to receive the charge packet from the well under $E2$ when a second clock pulse Φ_2 is applied. This completes a shift by one position.

13.3. COUNTERS

A counter is a circuit which counts the number of inputs (input pulses) it receives. Counters are widely used in various fields of radio engineering, applied electronics, and computers.

The performance of a counter is essentially stated in terms of its speed and count capacity (maximum count). The speed of a counter is a function of two parameters, namely the *maximum counting rate* f_c (or the minimum counting period, $T_c = 1/f_c$), and the *transition time*, τ_t, equal to the time interval between the arrival of the last count pulse and the display of the output count (word or code).

The count capacity of a counter is related to its *base* or *modulus*, μ. A modulo-μ counter has a count capacity of $\mu - 1$. Accordingly, a counter may be treated as a sequential circuit with μ internal states coded as a numeral N which can take on values 0, 1, 2, ... , $\mu - 1$. In our further discussion, the code N will be identified with the internal state of a counter. The code N is also the output count (word) of the counter.

The initial state is $N = 0$. As a rule, counters have a facility to clear (reset) them to zero. This operation is initiated by a reset (clear) pulse applied to the R (reset) input, and the counter cycles back to its initial state (zero).

Counters may be classed as *up-counters*, *down-counters*, and *up-down counters*. In an up-counter a 1 input increments the count by one; in a down-counter this input decrements the count by one. As its name implies, an up-down counter will increment or decrement the count, depending on the command signal received.

The above definitions are not rigorous enough to define the behaviour of a counter in an *overflow condition*, when the counter has reached its maximum (or minimum) count, and normal counting up or down cannot be carried on any longer. We shall consider counters with suitable outputs whose signals indicate when an overflow condition arises. Owing to these overflow outputs, as they are called, several counters can be cascaded so as to increase the overall count capacity. In rigorous terms, such counters can be specified, using the description of finite-state (sequential) machines.

Up-counter. This is a sequential circuit with a single data (count) binary input, T_{up}, and a binary carry output, P_{up}. It can be specifi-

ed by giving its transition function

$$N(t+1) = \begin{cases} N(t) & \text{for } T_{up}(t) = 0 \\ N(t)+1 & \text{for } T_{up}(t) = 1 \text{ and } N(t) \neq \mu - 1 \\ 0 & \text{for } T_{up}(t) = 1 \text{ and } N(t) = \mu - 1 \end{cases} \quad (13.13a)$$

and its output function

$$P_{up}(t) = \begin{cases} 0 & \text{for } T_{up}(t) = 0 \text{ or } N(t) \neq \mu - 1 \\ 1 & \text{for } T_{up}(t) = 1 \text{ and } N(t) = \mu - 1 \end{cases} \quad (13.13b)$$

From Eqs. (13.13) it follows that if at time t the count is $\mu - 1$, normal counting cannot be carried on any longer. In this state, a 1 input generates a signal $P_{up} = 1$ and resets (clears to zero) the counter (the counter cycles back to zero). The 1 at the P_{up} carry output is an indication of an overflow and is called an overflow one.

Down-counter. This is a sequential circuit with a single data (count) input, T_{down}, and a binary carry output, P_{down}. It can be specified by giving its transition and output functions

$$N(t+1) = \begin{cases} N(t) & \text{for } T_{down}(t) = 0 \\ N(t)-1 & \text{for } T_{down}(t) = 1 \text{ and } N(t) \neq 0 \\ \mu - 1 & \text{for } T_{down}(t) = 1 \text{ and } N(t) = 0 \end{cases} \quad (13.14a)$$

$$P_{down}(t) = \begin{cases} 0 & \text{for } T_{down}(t) = 0 \text{ or } N(t) \neq 0 \\ 1 & \text{for } T_{down}(t) = 1 \text{ and } N(t) = 0 \end{cases} \quad (13.14b)$$

If we compare the respective functions of an up-counter and a down-counter, it is seen that the only difference is that incrementing the count by one is replaced by decrementing the count by one. An overflow occurs when the down-counter reaches its minimum count, $N = 0$.

Up-down counter. Also known as a *forward-backward* or *reversible counter*, it can be made to count up or down (in the forward or the reverse direction), depending on the command signal used. It may have two count inputs or one count input and one sign input.

In the former case, it will have an up-count input T_{up}, an up-count carry output P_{up}, a down-count input T_{down}, and a down count carry output P_{down}. It will count up when $T_{down} = 0$, and down when $T_{up} = 0$. It is not usually permitted to apply a 1 signal to both count inputs at the same time. In the absence of inhibition, it is essential that the counter should not count when $T_{up}T_{down} = 1$, that is,

$$P_{up} = P_{down} = 0 \text{ and } N(t+1) = N(t) \quad (13.15)$$

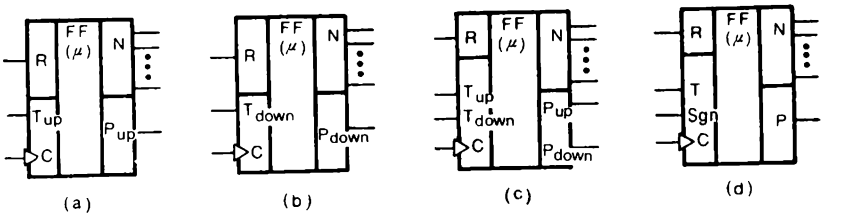


Fig. 13.23

In the latter case, a counter has a count input T, a carry output P, and a sign input Sgn which controls in which direction the counting should proceed. When Sgn = 0, it operates as an up-counter; when Sgn = 1, it operates as a down-counter.

Any of the above counters can be made synchronous and asynchronous (or ripple-through). Synchronous counters use synchronous (clocked) flip-flops (Fig. 13.23). In addition to its data inputs, each counter also has a clock input C which accepts the train of clock pulses maintaining the desired discrete time intervals in the circuit.

Asynchronous (ripple-through) counters are made with asynchronous flip-flops or with synchronous flip-flops but operated in the asynchronous mode when the sync input is used as a data input. The number and functions of the data inputs and outputs in an asynchronous counter are the same as in a synchronous counter. Externally, the operation of an asynchronous counter differs from that of a similar synchronous one in that the 1 signal ordinarily existing at the inputs T_{up}, T_{down} and T and the outputs P_{up}, P_{down} and P is replaced by a dynamic signal changing the values of the respective binary variables from 0 to 1 or from 1 to 0. In the former case, the asynchronous counter may be said to have a true dynamic input, and in the latter, a complementing dynamic input. Figure 13.24 shows symbols for several types of synchronous counters. The Sgn input of an up-down counter is designated as a static input, because by itself it cannot change the state of the counter (a control input).

Counters in cascade. A cascade connection of counters is used in order to increase the count capacity in excess of what the individual counters can give. To form a cascade, the carry output of one

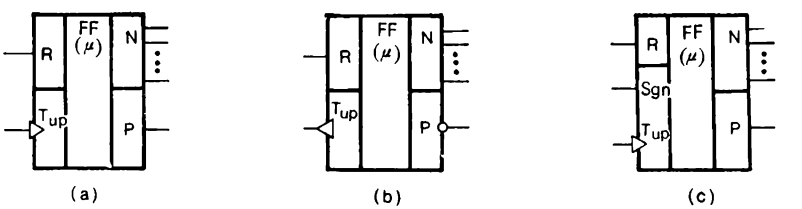


Fig. 13.24

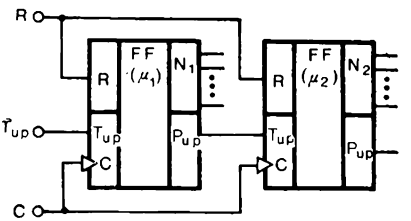


Fig. 13.25

counter is taken to a similar count input of the another, say P_{up} to T_{up} or P_{down} to T_{down}, etc. As an example, consider the operation of the counter shown in Fig. 13.25. It is a cascade of a modulo-μ_1 and a modulo-μ_2 counter. Suppose that initially both counters are in the initial (zero) state. After the count input of the first counter has received μ_1 units, an overflow condition arises. The carry signal reaching the input of the second counter is recognized as a count pulse. As a result, the first counter cycles back to 0, and the count N_2 on the second is incremented by one. This situation will occur again after the input of the first counter has received another μ_1 units. The total number of units fed to the counter input will be

$$N_\Sigma = \mu_1 N_2 + N_1 \tag{13.16}$$

Both counters generate an overflow signal after they have received $\mu_1\mu_2$ units. Thus, two counters in cascade count a modulo which is the product of their respective moduli. A similar result can be obtained, using a cascade of down-counter and up-down counters. It can readily be extended to a cascade of an arbitrary number of counters. In a typical situation, all the counters in cascade have the same modulus, $\mu_1 = \mu_2 = \ldots = \mu_n = \mu$. Then the total number of units applied to the counter input is

$$N_\Sigma = \mu^{n-1}N_n + \mu^{n-2}N_{n-1} + \ldots + \mu^0 N_1 \tag{13.17}$$

which is a generalization of Eq. (13.16).

Equation (13.17) shows that a cascade of counters with the same modulus may be treated as a counter whose output count is a code in a base μ number system.

Binary Counters

A binary counter is a counter which counts in the binary number system: its modulus is an integer power of 2 and the output count is a binary number:

$$N = 2^{n-1}N_n + 2^{n-2}N_{n-1} + \ldots + 2^0 N_1 \tag{13.18}$$

where $N_1, N_2, \ldots, N_n$ are binary variables.

From a comparison of Eqs. (13.18) and (13.17) it is seen that a binary counter can be built by connecting in cascade a number n of cells or stages each of which is a modulo-2 counter ($\mu = 2$). In our further discussion we shall call this elementary counter a binary cell or, simply, a binary.

A binary is a sequential circuit with two internal states and usually contains one flip-flop whose binary output variable Q gives the output count, that is, $N = Q$. In addition to a flip-flop, a binary may include logic circuits to produce steering and carry signals.

The transition and output functions of a binary can readily be derived from the general equations (13.12) through (13.14) which, upon replacement of N by Q, are written as switching functions:

—for an up-count binary:

$$Q(t+1) = Q(t) \oplus T_{\text{up}}(t), \quad P_{\text{up}}(t) = Q(t)\, T_{\text{up}}(t) \tag{13.19a}$$

—for a down-count binary:

$$\begin{aligned} Q(t+1) &= Q(t) \oplus T_{\text{down}}(t) \\ P_{\text{down}}(t) &= \overline{Q}(t)\, T_{\text{down}}(t) \end{aligned} \tag{13.19b}$$

—for an up-down binary with two count inputs:

$$\begin{aligned} Q(t+1) &= Q(t) \oplus [T_{\text{down}}(t) + T_{\text{up}}(t)] \\ P_{\text{up}}(t) &= Q(t)\, T_{\text{up}}(t)\, \overline{T}_{\text{down}}(t) \\ P_{\text{down}}(t) &= \overline{Q}(t)\, \overline{T}_{\text{up}}(t)\, T_{\text{down}}(t) \end{aligned} \tag{13.19c}$$

—for an up-down binary with a sign input:

$$\begin{aligned} Q(t+1) &= Q(t) \oplus T(t) \\ P(t) &= [Q(t) \oplus \text{Sgn}(t)]\, T(t) \end{aligned} \tag{13.19d}$$

From a comparison of Eqs. (13.19a) and (13.19b) it is seen that to change from up-counting to down-counting it will suffice to invert the variable in the equivalence defining the function of output carry. In this way, one can readily convert up-counters into down-counters by replacing the true output of the flip-flop in the carry circuit by a complementing one or, conversely, by making the complementing output a true one.

Synchronous binaries and counters. This class of binaries is usually made with synchronous T flip-flops. The structure of the binaries can be configured immediately from Eqs. (13.19). The circuit of an up-count binary is shown in Fig. 13.26*a*, and the transition function (13.19) is implemented by applying the count pulse to the T-input of the flip-flop. To generate the carry signal, the circuit is extended to include an AND gate. The flip-flop may be a JK or a DV type. Figure 13.26*b* shows an up-count binary realized with a DV

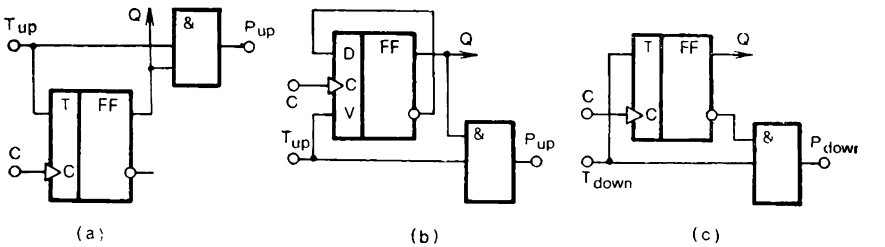


Fig. 13.26

flip-flop. Similarly, the T flip-flop can be replaced by a universal flip-flop in all the circuits examined below.

A down-count binary (Fig. 13.26*c*) differs from an up-count binary in that the variable Q at the input of the AND gate is replaced by its complement.

Up-down binaries with two count inputs are shown in Fig. 13.27. The structure of the circuit in Fig. 13.27*a* is uniquely specified by Eqs. (13.19*c*). The circuit in Fig. 13.27*b* is derived on the assumption that it is not permitted to apply 1's to the T_{up} and T_{down} inputs at the same time. The terms $T_{up}\overline{T}_{down}$, $\overline{T}_{up}T_{down}$, and $T_{down} \oplus T_{up}$ replace, respectively, T_{up}, T_{down}, and $T_{up} + T_{down}$. In designing the circuit in Fig. 13.27*b*, the basic Boolean functions have been replaced by NAND functions (see Chap. 5).

An up-down binary with a sign input, answering Eqs. (13.19d), is shown in Fig. 13.28. When Sgn = 0, it operates as the circuit in Fig. 13.16*a*; when Sgn = 1, it operates as the circuit in Fig. 13.26*c*.

The mean counting rate of a synchronous counter is a maximum when a 1 signal is applied to the data input during each clock time. Then the speed of the counter is determined by the minimum clock pulse repetition period just sufficient for normal operation. Because all transients in the flip-flops and the associated gates must completely die out by the instant when the next pulse arrives, the minimum

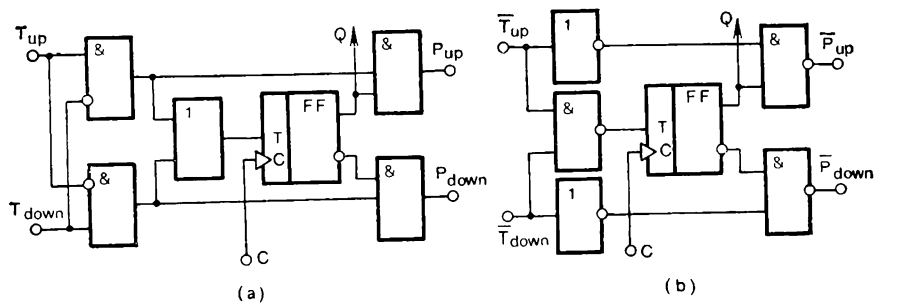


Fig. 13.27

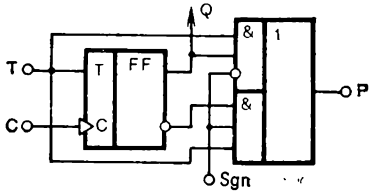


Fig. 13.28

time between pulses is $T_{c,\,min} = \tau_F + \tau_{\Sigma G}$, where τ_F is the propagation delay of one flip-flop and the $\tau_{\Sigma G}$ is the propagation delay through all the associated gates.

The value of $\tau_{\Sigma G}$ depends on the counter structure. For binary counters examined so far, which are binaries in cascade, the speed of operation is limited by the carry propagation delay.

Let τ_P denote the carry propagation delay of a binary relative to the instant when a count pulse T_{up} is applied. Taking the latter instant as the origin of time, the count pulse will appear at the input of the second binary after a time delay equal to τ_P, at the input of the third binary after a delay $2\tau_P$, etc. The time delay will be a maximum equal to $(n-1)\,\tau_P$ at the input to the last (n th) binary containing the most significant bit of the output word. This carry propagation delay determines the value of $\tau_{\Sigma G}$. Hence, the minimum repetition period for synchronous counters made with several binaries in cascade is

$$T_{c,\,min} = \tau_F + (n-1)\,\tau_P \tag{13.20}$$

Equation (13.20) shows that with synchronous counters in cascade the maximum counting rate decreases.

The time required for the output word to form (the transition time) in the synchronous counters examined above is the same as the clock pulse repetition period and is not a decisive parameter.

Asynchronous binaries and counters. In asynchronous form, the transition and output functions can be derived from Eqs. (13.19) on replacing the level-type variables T_{up}, T_{down}, T, P_{up}, P_{down} and P by their dynamic analogues, namely $T_{up}\mathrm{d}T_{up}$, $T_{down}\mathrm{d}T_{down}$, $T\mathrm{d}T$, $P_{up}\mathrm{d}P_{up}$, $P_{down}\mathrm{d}P_{down}$ and $P\mathrm{d}P$ if the circuit is triggered by the negative edge of the signal, and $\overline{T}_{up}\mathrm{d}T_{up}$, $\overline{T}_{down}\,\mathrm{d}T_{down}$, $\overline{T}\mathrm{d}T$, $\overline{P}_{up}\mathrm{d}P_{up}$, $\overline{P}_{down}\mathrm{d}\overline{P}_{down}$, and $\overline{P}\mathrm{d}P$, if the binary is triggered by the positive edge of the signal.

For a positive-edge-triggered up-count binary, from Eq. (13.19a) we get

$$Q(t+1) = Q(t) \oplus \overline{T}_{up}\,\mathrm{d}T_{up} \tag{13.21a}$$

$$\overline{P}_{up}\,\mathrm{d}P_{up} = Q(t)\,\overline{T}_{up}\,\mathrm{d}T_{up} \tag{13.21b}$$

The transition functions defined by Eq. (13.21b) are implemented by an asynchronous T flip-flop. In differential form, this function may be written

$$\mathrm{d}Q(t) = \overline{T}_{up}(t)\,\mathrm{d}T_{up} \tag{13.22}$$

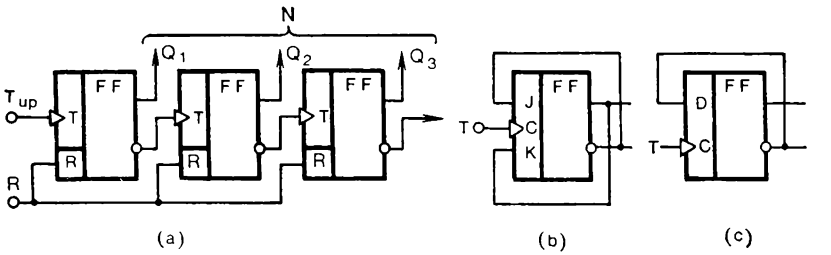


Fig. 13.29

On substituting Eq. (13.22) into (13.21a), the transition function may be written

$$\overline{P}_{up}\, dP_{up} = Q(t)\, dQ(t) = \overline{\overline{Q}}(t)\, d\overline{Q}(t) \qquad (13.23)$$

As is seen, the carry signal may be the positive step change in the signal existing at the complementing output of the T flip-flop, that is, $P_{up} = \overline{Q}$.

Three stages of an asynchronous (ripple) counter based on asynchronous positive-edge-triggered T flip-flops are shown in Fig. 13.29. Figures 13.29*b* and *c* show how the terminals of clocked JK and D flip-flops should be connected so as to obtain the counting mode, using the C input as the data input.

As already noted, an asynchronous up-count binary can readily be converted into a down-count binary, if we replace the complementing input on the flip-flop in the carry line by a true input.

For negative-edge triggering, it is necessary either to place an inverter at the counter input or to use flip-flops having a complementing dynamic count input. On replacing $\overline{P}_{up}$ by P_{up} in Eq. (13.23), we obtain

$$\overline{P}_{up}\, dP_{up} = Q(t)\, dQ(t)$$

That is, the carry signal for a negative-edge-triggered up-count binary should be the true output of the flip-flop (Fig. 13.30). Conversely, for a down-count binary, this should be the complementing output.

In the asynchronous binaries and counters examined above, the carry propagation delay is related to the switching time of the flip-flops. In some cases, this delay can be minimized, using what might be called "ripple-carry" circuits where the variable P is generated as a switching function of the input signal and the state of the flip-flop.

If we take negative-edge-triggered binaries, then for an up-count binary the switching function of the carry operation will be

$$P_{up} = T_{up} Q \qquad (13.24a)$$

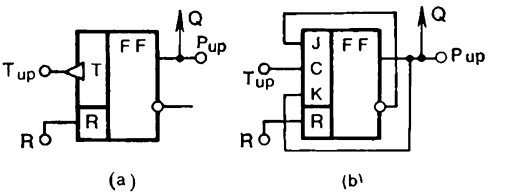


Fig. 13.30

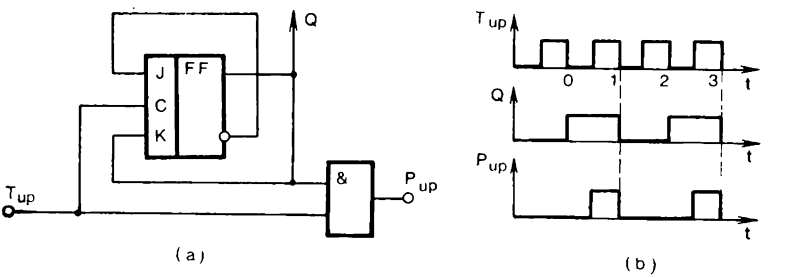


Fig. 13.31

and for a down-count binary,

$$P_{\text{down}} = T_{\text{down}}\overline{Q} \tag{13.24b}$$

The circuit of a clocked JK flip-flop up-count binary and its timing diagrams are shown in Fig. 13.31*a* and *b*, respectively.

The "carry" switching functions for positive-edge-triggered binaries can be derived from Eq. (13.24) on replacing the variables P_{up}, P_{down}, T_{up}, and T_{down} by their complements. Then, according to De Morgan's theorems

$$P_{\text{up}} = T_{\text{up}} + \overline{Q}, \quad P_{\text{down}} = T_{\text{down}} + Q \tag{13.25}$$

Asynchronous up-down counters with two count inputs utilize ripple carry. Figure 13.32 shows a stage of a such a counter triggered by the negative edge of the input signal. It is not permitted to apply $T_{\text{up}} = 0$ and $T_{\text{down}} = 0$ at the same time.

An asynchronous counter with a sign input can be built using the circuit in Fig. 13.33. It converts the signals Sgn and T into a pair of count signals. T_{up} and T_{down}. When it is connected to the input of a counter with two count inputs, it will supply pulses to the up-count input when Sgn = 0 and to the down-count input when Sgn = 1.

Each binary may be treated as a PRF (PRR) divider. The maximum counting rate of a complete counter is limited by the maximum switching rate of the flip-flop in the first binary containing the least

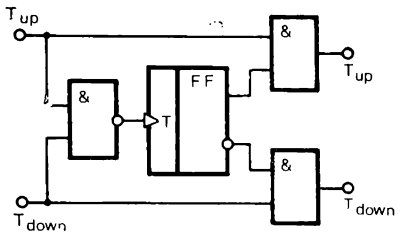


Fig. 13.32

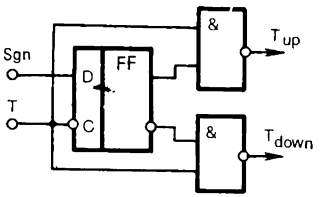


Fig. 13.33

significant bit of the output word. The stages storing the more significant bits may be circuits having a lower speed as compared with the stages storing the less significant bits.

The transition time τ_t of asynchronous counters depends on the carry propagation delay of the chain of the binaries and on the final count. It is a maximum when the binary variables change in value all at the same time and in all the stages. In up-counting, this situation arises when $N = 0$ and $N = 2^{n-1}$, and also in the case of an overflow. The maximum transition time is given by

$$\tau_{t,\,\max} = \tau_F + (n-1)\,\tau_P$$

where τ_F is the propagation delay of one flip-flop and τ_P is the carry propagation delay of one binary.

In counters without ripple carry, $\tau_P = \tau_F$, whereas in counters with ripple carry τ_P depends on the propagation delay through the logic gate and is usually smaller than τ_F.

The transition time is a minimum when the binary variable is changed in only one of the stages. In up-counting, this situation arises when $N = 2^m + 1$, where $m = 0, 1, \ldots, n-1$. The minimum transition time depends on the speed of the flip-flop storing the least significant bit: $\tau_{t,\,\min} = \tau_F$.

Parallel-carry binary counters. In contrast to ordinary synchronous counters (or asynchronous counters with ripple carry), these counters use a parallel rather than a cascade connection of the "carry" gates.

For insight into the principle underlying the design of parallel (or look-ahead) carry binary counters, we shall examine up-counters analogous to those shown in Figs. 13.26*a* and Fig. 13.31*a*. In either case, the carry signal is generated by ANDing the input signal T_{up} and the state of the flip-flop. This can be described in terms of a recursive equation

$$T_{\text{up},\,k} = T_{\text{up},\,k-1} Q_{k-1}$$

where $T_{\text{up},\,k}$ is the value of the binary variable at the data input of the kth binary, and Q_k is the state of the flip-flop used in the kth

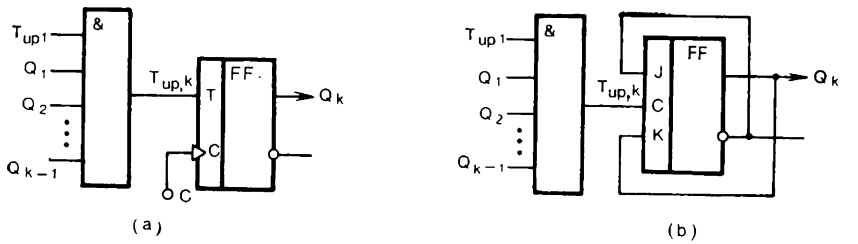


Fig. 13.34

binary. On expanding the above equation consecutively, we get

$$T_{\text{up},k} = T_{\text{up},k-1}Q_{k-1} = T_{\text{up},k-2}Q_{k-2}Q_{k-1}$$
$$= \ldots = T_{\text{up},1}Q_1 \ldots Q_{k-1} \tag{13.26}$$

The design of a parallel-carry counter reduces to implementing the right-hand side of Eq. (13.26) separately for each stage of the counter. The synchronous and asynchronous binaries thus developed appear in Fig. 13.34*a* and *b*, respectively.

The minimum pulse period for a synchronous parallel-carry counter (see Fig. 13.34*a*) is

$$T_{c,\min} = \tau_F + \tau_G \tag{13.27}$$

where τ_G is the propagation delay of the input AND gate.

For an asynchronous parallel-carry counter (see Fig. 13.34*b*) the transition time is

$$\tau_t = \tau_F + \tau_G \tag{13.28}$$

From Eqs. (13.27) and (13.28) it is seen that the speed of parallel-carry counters does not depend on the number of stages used.

In some cases, the AND function can be implemented, using the internal logic gates of the flip-flop. In such a case, the speed of parallel-carry counters will solely depend on the propagation delay through the flip-flop proper.

The need to have a larger number of inputs on binary cells for the more significant bits is a disadvantage of parallel-carry counters. Sometimes, a combination of parallel-carry counters in cascade may fill this need.

13.4. COMBINATIONAL DIGITAL FUNCTIONAL SUBUNITS

The basic digital microoperations performed by combinational logic networks are addition, comparison, and code conversion.

Adders

An adder is a combinational logic circuit which sums two numbers. Digital systems mostly use binary adders because of their simplicity and economy of components.

Two multibit numbers may be added serially (one column at a time) or in parallel (all columns simultaneously). Obviously *serial addition* can be done by one and the same one-bit adder used repeatedly. *Parallel addition* requires as many one-bit adders as there are the bits being summed, and is carried out within a single bit time.

A one-bit adder may have three inputs: the addend a, the augend b, and the carry c from the next lower bit, or the input carry (Fig. 13.35*a*); this is a *full adder*. Or it may have only two inputs: the addend a and the augend b, but no input for the carry from the next less significant bit (Fig. 13.35*b*); this is a *half adder*.

The addition of two or three one-bit numbers gives a two-bit number. The least significant bit is the sum, S, and the most significant bit is the carry, C, to the next higher bit (of another adder). Accordingly, there are two outputs, S for the sum bit and C for the output carry.

In Tables 13.1 and 13.2, the output binary variables S and C are defined as the switching functions of the input variables a and b (c). The respective switching (Boolean) expressions are:

Table 13.1. Switching Functions of a Full Adder

a	b	c	C	S	a	b	c	C	S
0	0	0	0	0	1	0	0	0	1
0	0	1	0	1	1	0	1	1	0
0	1	0	0	1	1	1	0	1	0
0	1	1	1	0	1	1	1	1	1

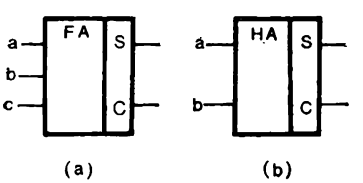


Fig. 13.35

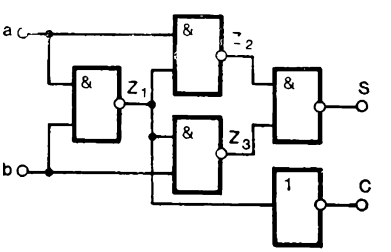


Fig. 13.36

Table 13.2. Switching Functions of a Half Adder

a	b	C	S
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

—for a full adder

$$S = abc + a\bar{b}\bar{c} + \bar{a}b\bar{c} + \bar{a}\bar{b}c$$
$$C = ab + ac + bc \tag{13.29}$$

—for a half adder

$$S = \bar{a}b + a\bar{b},\ C = ab \tag{13.30}$$

The likely logic diagrams of a half-adder and a full adder are shown in Figs. 13.36 and 13.37. It is left as an exercise for the reader to verify the correspondence of these diagrams to Eq. (13.29) or Eq. (13.30) by writing a set of structural equations (see Sec. 5.1), substitution and identity transformations in Boolean algebra. In these diagrams, all inputs (a, b and c) and all outputs (S and C) are true. In some cases, adders may have inverting (complementing) inputs and/or outputs. This will usually simplify the adder circuit and enhance its speed. An example is the full adder arranged as shown in Fig. 13.37*b*, with the output inverters omitted. The invert-

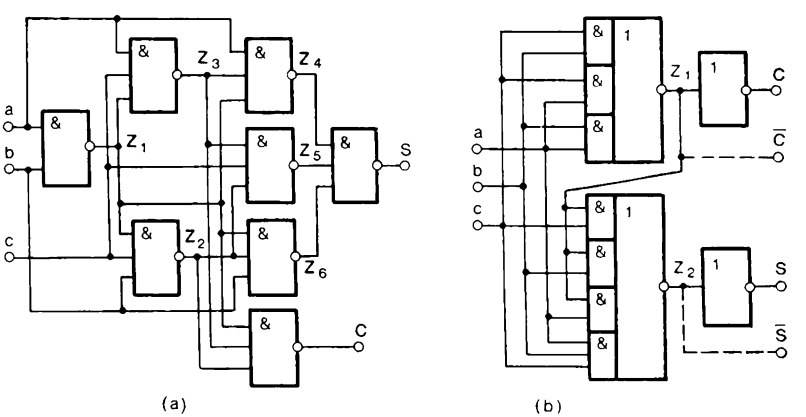


Fig. 13.37

ed outputs $\bar{S}$ and $\bar{C}$ are generated directly by AND-NOR gates (the dashed lines in Fig. 13.37*b*). The graphic symbol for a full adder with inverted outputs is shown in Fig. 13.38*a*.

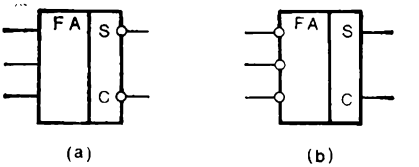


Fig. 13.38

A distinction of the functions $S\ (a,\ b,\ c)$ and $C\ (a,\ b,\ c)$ implemented by a full adder is their self-duality*. This implies that the two functions are such that

$$S\ (a,\ b,\ c) = \bar{S}\ (\bar{a},\ \bar{b},\ \bar{c}),\quad C\ (a,\ b,\ c) = \bar{C}\ (\bar{a},\ \bar{b},\ \bar{c})$$

The self-duality of the switching functions of a full adder implies that the inversion of the binary variables at all of its inputs results in the inversion of the output variables, irrespective of their values. Therefore, an adder with inverted outputs (Fig. 13.38*a*) may be

* The function $f^*\ (x_1,\ \ldots,\ x_n)$ is a dual of the function $f\ (x_1,\ \ldots,\ x_n)$, if $f^*\ (x_1,\ \ldots,\ x_n) = \bar{f}\ (\bar{x}_1,\ \ldots,\ \bar{x}_n)$. A function is said to be self-dual, if it is identical with a function which is its own dual.

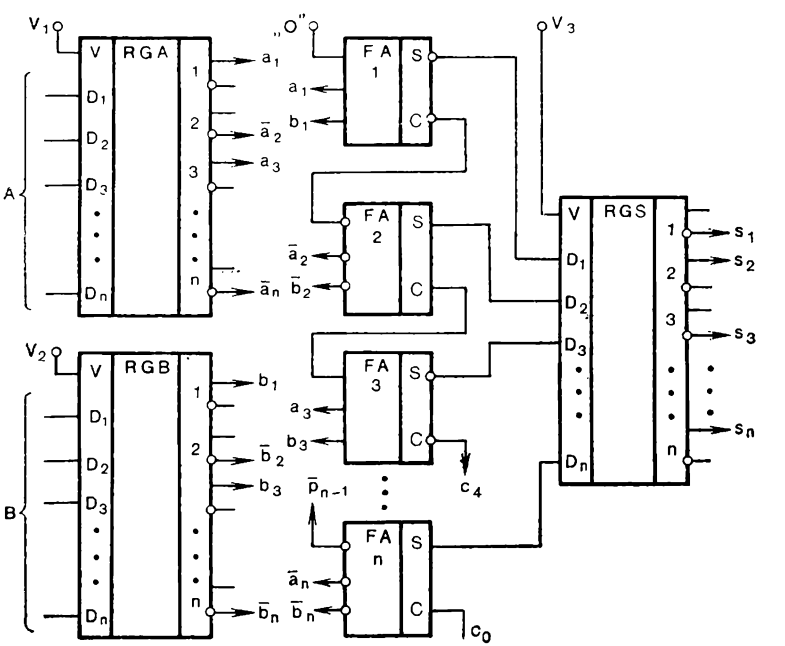


Fig. 13.39

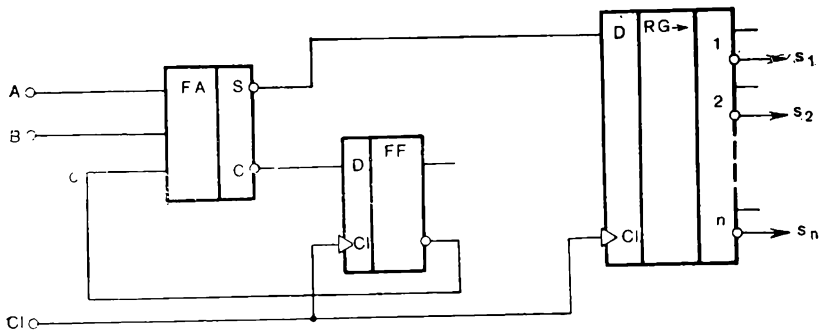


Fig. 13.40

interpreted as being equivalent to an adder with inverting inputs such as shown in Fig. 13.38*b*.

Adders with inverting inputs (outputs) can conveniently be used in conjunction with symmetrical flip-flops which, in addition to the memory function, may also perform inversion.

As typical examples of multibit adders, we shall discuss the circuits shown in Figs. 13.39 and 13.40.

Parallel adder circuit. This type of adder (Fig. 13.39) has an addend register RGA, an augend register RGB, and a sum register RGS. The combinational part consists of n full adders with signal inversion. The carry line is formed by connecting the C output of the ith stage to the input of the $(i + 1)$st stage. Each odd adder is treated as a full adder with inverting outputs, and each even adder as a full adder with inverting inputs. Also, a true output is always taken to a true input, and an inverting output to an inverting input.

Serial adder circuit. A serial adder (Fig. 13.40) differs from a parallel adder in that two multibit numbers are added serially, that is, one column at a time, starting with the least significant bits. Accordingly, the addend and the augend are fed to a full adder serially. The carry signal C is generated during the previous bit time and delayed for one bit time by a D flip-flop. The result is written into a serial register.

Digital Equality Detectors

A *digital (or binary) equality detector* serves to check the equality of two binary numbers. Two n-bit numbers X and Y are equal, if a bit in a particular position of one number is equal to the bit of the same significance in the second number, that is, if the equality $a_i b_i + \bar{a}_i \bar{b}_i = 1$ holds for all the i bits. Accordingly, the switching function Y implemented by a digital equality detector may be writ-

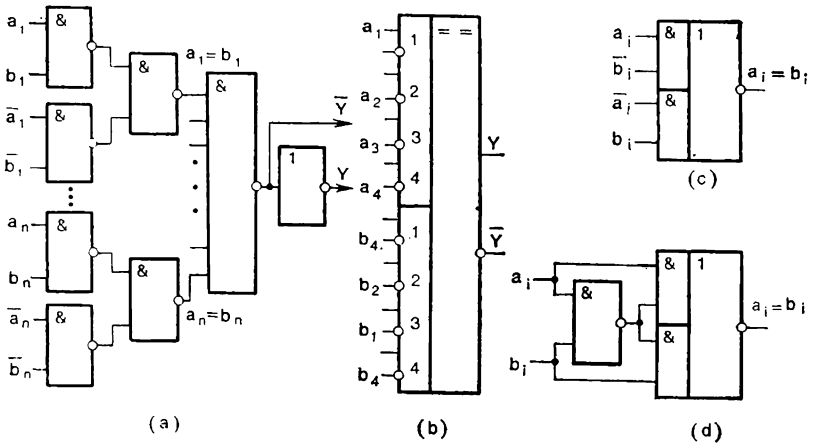


Fig. 13.41

ten

$$Y = \prod_{i=1}^{n} (a_i b_i + \bar{a}_i \bar{b}_i)$$

Figure 13.41*a* shows a logic block diagram of a digital equality detector made with NAND gates; its logic symbol appears in Fig. 13.41*b*. The likely modifications of an equality detector stage based on AND-NOR gates are shown in Fig. 13.41*c* and *d*. Since the bits of the numbers being compared are handled all at the same time, the total delay of the circuit only slightly depends on the number of bits, and ranges from $3\bar{\tau}_d$ to $5\bar{\tau}_d$.

Decoders/Demultiplexers

A *decoder* (*demultiplexer*) is a functional subunit with n inputs which excites one and only one output for each of the likely combinations of input signals (input codes). If a decoder uses all the likely binary combinations, the number W of outputs is $W = 2^n$. This will be called a complete decoder in contrast to a partial decoder which uses only some of the likely input combinations. We shall only be concerned with complete decoders.

A need for a decoder (demultiplexer) arises when a binary signal (serial data) is to be routed into one of W lines, the desired line being designated by an n-bit code (address). If we arrange the inputs and outputs of a decoder in increasing order of bit significance, so that the input code $x_1, x_2, x_4, \ldots, x_{2^{n-1}}$ causes the excitation of an output whose number is $w = x_1 + 2x_2 + 4x_4 + \ldots + 2^{n-1}x_{2^{n-1}}$, then W ranged expressions can be written for the outputs of a com-

plete decoder. Let, for example, $n = 3$, and $W = 8$. Then,

$$Y_0 = \bar{x}_4\bar{x}_2\bar{x}_1$$
$$Y_1 = \bar{x}_4\bar{x}_2x_1$$
$$Y_2 = \bar{x}_4x_2\bar{x}_1$$
$$\cdots\cdots$$
$$Y_7 = x_4x_2x_1$$

Linear decoder. When the above switching functions are implemented with AND gates, the result is a linear decoder (demultiplexer), the simplest of all as regards structure.

Consider some aspects of how such decoders can be made with NAND gates.

When a decoder is built with NAND gates, its outputs are not the functions Y_i, but their complements. In other words, only one selected output is in a 0 state, whereas all the remaining (unselected) outputs are in a 1 state. Because of this, the outputs on the logic symbol of such a decoder must be denoted by inversion symbols (Fig. 13.42). In order to implement true outputs, the decoder must be supplemented by an inverter array (which is implied in Fig.13.42*a*). The inverted output of the decoder matches well the set and reset inputs of NAND gate flip-flops (Fig. 13.42*b*). Now the *RS* flip-flop is set to a 1 state when the zero output of the decoder is excited. When the output of a decoder is coupled to clocked *D* flip-flops, it is likewise advisable to use an inverted signal because the inversion can be performed by the flip-flop itself.

The expressions describing the input-output behaviour of a decoder involve not only the input variables but also their complements. Accordingly, the inputs of the NAND gates must be driven by both the true and the complement values of the input variables. To satisfy this, either true/complement connections must be used from the decoder to the register or the signals must be inverted inside the decoder by an inverter. In a linear decoder, each input is loaded

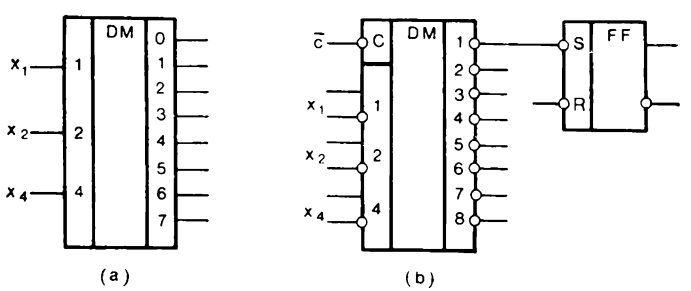


Fig. 13.42

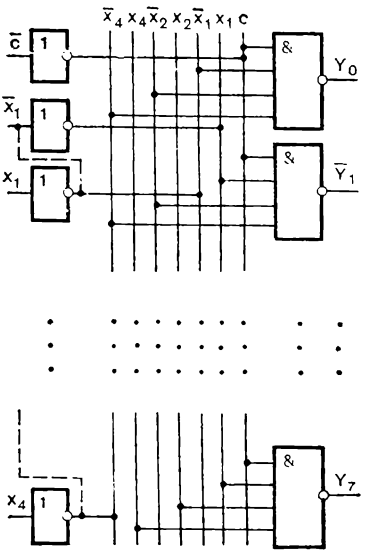


Fig. 13.43

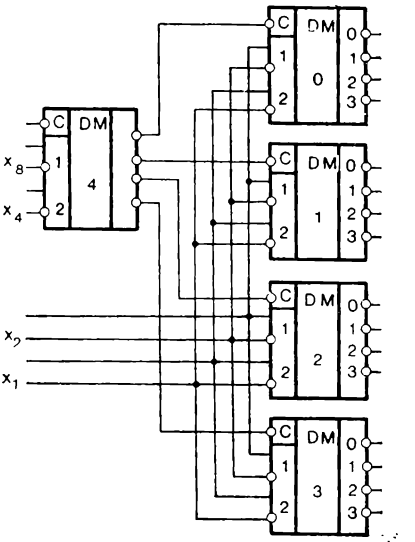


Fig. 13.44

by $W/2$ NAND gates. Since the register may be loaded into several functional subunits, it is advisable to provide at the decoder input several distributors (inverters, Fig. 13.43) which permit a sizeable decrease in the load that the decoder puts on the register flip-flops.

A demultiplexer is in effect a multi-output combinational network, and as in any such network the signals propagate from input to output with a different time delay. Quite logically, this is fraught with the race problem. The race condition can be avoided by "sampling" (or gating) the otuput signal, which can be done by clocked flip-flops. A more economical approach is to combine decoding and gating in the same NAND gates. For this purpose each NAND gate must have one input more than the length n of the code to be handled, and a gating (strobe) signal c must be applied to this additional input as shown in the circuit of Fig. 13.43. Since the gating signal is fed to a large number of logic gates, it is usual to provide at the demultiplexer input one more distributor-inverter which inverts this input (Fig. 13.42*b*).

High-order demultiplexers. Figure 13.44 shows a group of five demultiplexers arranged in two-section "tree" formation. All the demultiplexers are identical and have structure analogous to that shown in Fig. 13.43. The first section, or trunk four-output demultiplexer, $DM4$, accepts the more significant bits of the input number, x_8 and x_4. $DM4$ determines which four out of the likely 16 combinations

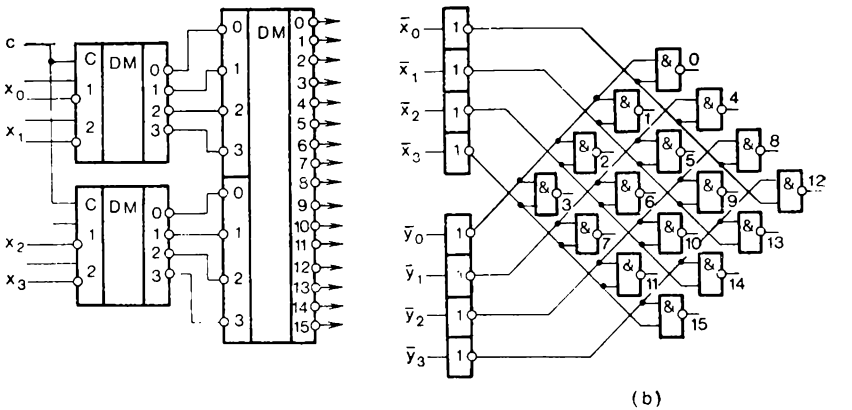


Fig. 13.45

contain the desired output. The trunk enables the corresponding branch, or demultiplexer in the second section, by feeding a 0 signal to the c input, because all the other outputs of $DM4$ generate a 1 signal; the remaining "branch" demultiplexers remain disabled. The less significant bits of the input code go directly to the branches where only one demultiplexer is operating, so that out of the 16 output lines only one corresponding to the input code is excited. For example, if the code applied to inputs x_8, x_4, x_2, and x_1 is 1011, only output *2* will be excited in $DM4$, so that only $DM2$ will be enabled, and the other demultiplexers in the second section ("branches") will remain disabled. The least significant bits of the input code, $x_2 = 1$ and $x_1 = 1$, decoded in $DM2$, will excite its output *3*. As a result, output 11 will be selected, which corresponds to the input code.

Thus, high-order demultiplexers with a large number of outputs can be built using low-order demultiplexers as building blocks. Low-order demultiplexers are usually manufactured as standard ICs or plug-ins.

Because the most essential components of a high-order demultiplexer are concentrated in its last section, it will be advantageous to make its elements as simple as practicable. An ideal case would be if each output signal were generated by a two-input gate. This would however lead to an increased number of gates in the preceding section. By far the best solution is offered by rectangular demultiplexers.

Rectangular demultiplexer (Fig. 13.45a). This network is made up of two demultiplexers, $DM1$ and $DM2$, in the first section and one matrix demultiplexer, $DM3$ in the second (Fig. 13.45*b*). The high- and the low-order bits are decoded independently by $DM1$

and $DM2$, respectively. As a result, the output of the matrix demultiplexer produces a signal generated by a NAND gate at the intersection of the excited input lines. If, for example, a 0 signal be applied to inputs y_2 and z_1, the output signal will appear on output *6*.

The second section of a rectangular demultiplexer contains a number W of two-input NAND gates. With a given W, the size of $DM1$ and $DM2$ will depend on how the combination of input variables is partitioned. Let $DM1$ have W_1 outputs and $DM2$, W_2 outputs. Then $W_1W_2 = W$. and $W_1 + W_2$ indicates how many gates are required for the first section. For a given product, the sum of the terms is a minimum when these terms are equal.So the input variables should preferably be partitioned into two equal groups. If n is even, then $W_1 = W_2 = \sqrt{W}$. This dependence indicates that the share of first-section components in the total equipment count will decrease with increasing W, so with high values of W the bulk of components will be concentrated in the second section.

13.5. ANALOG COMPARATORS

An *analog comparator* (or *detector*) is a switching network converting the sign of the difference between two analog variables x_1 and x_2 into a binary output variable y. In fact, an analog comparator acts as an analog-to-digital converter.

Assuming that y is represented as a potential (voltage) level, the transfer characteristic of an ideal analog comparator may be depicted as shown in Fig. 13.46*a*. It implies that when $x_1 > x_2$, the comparator output goes high which corresponds to $y = 1$. When $x_1 < x_2$, the output goes low, which corresponds to $y = 0$. The logic levels V^H and V^L must be matched to the logic gates used in the digital portion of the comparator.

Real switching circuits will have transfer characteristics (Fig. 13.46*b*) similar to those of an ideal analog comparator only if the difference signal, $x = x_1 - x_2$, is sufficiently large in absolute value. At low values of $|x|$, the value of the signal at the output of an analog comparator cannot be predicted with certainty because of many factors, such as drift, fluctuation noise, variations in supply

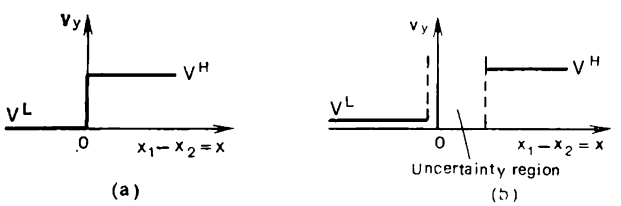


Fig. 13.46

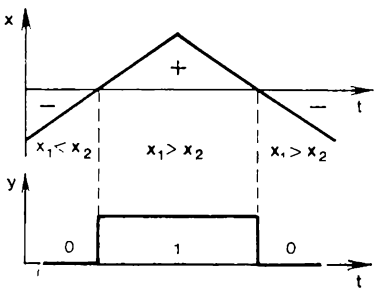


Fig. 13.47

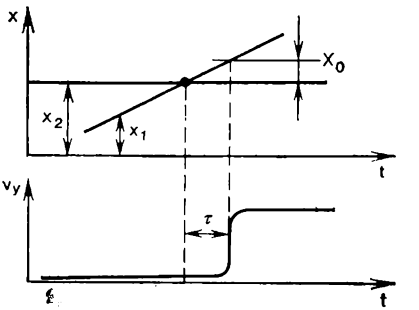


Fig. 13.48

voltages, and hysteresis. The width of the uncertainty region is closely related to a major characteristic of analog comparators—sensitivity defined as the minimum value of the input signal just sufficient for the output variable to go from 1 to 0 or back reliably. Within the uncertainty region, the behaviour of an analog comparator will differ from type to type.

Analog comparators may be classed into asynchronous and synchronous, according as their output variable y varies with a change in the input variables.

Asynchronous analog comparators. An ideal asynchronous comparator generatesat its output a step change of voltage just as the difference signal $x\,(t)$ passes through zero. Such a configuration is referred to as a zero-crossing detector. If the difference signal is positive-going, a positive step change, $\bar{y}dy$, is formed. If the difference signal is negative-going, a negative step change, ydy, is produced (Fig. 13.47).

Real asynchronous comparators are often intended to respond to a change in the difference signal in one direction only—either positive-going or negative-going. In the general case, they generate an output signal at a time other than when $x\,(t) = 0$, the difference between the two instants being τ (Fig. 13.48). The accuracy of an analog comparator is stated in terms of the error referred to its input, also called the comparison error defined as the value of the difference signal $x = X_0$ at the instant when the output step change is formed.

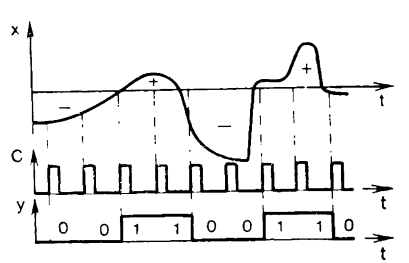


Fig. 13.49

Synchronous analog comparators. These networks differ from the asynchronous type in that the output variable changes its value

only in response to a clock pulse and is in no way related to the instant when the difference signal changes its sign (Fig. 13.49). If we reference the discrete (sampling) time intervals to the leading edge of the clock pulse C, the transfer function of an ideal synchronous analog comparator may be written

$$y(t+1)=\begin{cases} 1 & \text{for} \quad x_1(t) > x_2(t) \\ 0 & \text{for} \quad x_1(t) < x_2(t) \end{cases}$$

As already noted, real analog comparators differ from ideal ones in that they have an uncertainty region. The operation of a synchronous analog comparator (both within and outside its uncertainty region) can rigorously be described, if we treat the appearance of a 1 or a 0 at its output as two mutually complementing or exclusive random events.

The systematic comparison error $\bar{X}_0$ of a synchronous analog comparator is defined as the static value of the difference signal such that a 0 and a 1 have an equal probability of appearing at the output.

The random component of the comparator error, ΔX_0, is defined as the minimum absolute deviation of the difference signal from $\bar{X}_0$, such that the comparator will operate reliably with a specified confidence level at time t^* and an arbitrary $y(t^* - 1)$. Then the points $\bar{X}_0 + \Delta X_0$ and $\bar{X}_0 - \Delta X_0$ may be regarded as the upper and lower boundaries of the uncertainty region on the static characteristic of the comparator.

In contrast to asynchronous analog comparators, no restraint is put on the rate of change in the difference signal between clock pulses in synchronous comparators. Therefore, quantitatively the dynamic error of a synchronous comparator can be found by analyzing its response to a standard step change in the difference signal, X_m (for voltage comparators, X_m is usually taken equal to 100 mV). The position of the step change (Fig. 13.50*a*) depends on the time by which it leads the clock pulse, and the overexcitation ΔX equal

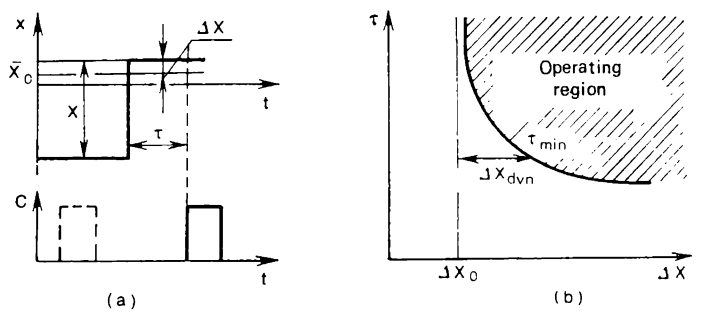


Fig. 13.50

to the amount by which the input signal exceeds the threshold signal $\bar{X}_0$. In the general case, whether the response of a comparator is true or false depends on the relation between τ and ΔX. For any $\Delta X > \Delta X_0$, there is a finite value τ_{min} (Fig. 13.50*b*) such that the comparator will operate with the specified confidence level.

The dynamic error, ΔX_{dyn} may be defined as the amount by which the overexcitation exceeds the random error ΔX_0 at a given τ_{min}. The dynamic error tends to zero as τ tends to infinity, so the required overexcitation is completely determined by the random component of the comparator error.

The two analog quantities being compared may be different physically. Accordingly, there may be amplitude, phase, and frequency comparators. Amplitude comparators compare the instantaneous values of voltages (or currents), whereas phase and frequency comparators compare the characteristics of quasi-periodic pulse signals.

Amplitude Analog Comparators

An amplitude analog comparator (Fig. 13.51*a*) is a difference detector, *DD*, and a threshold circuit, *TC*, connected in cascade. The difference detector generates an output voltage, v_d, which is proportional to the difference between the quantities being compared:

$$v_d = K(x_1 - x_2) \tag{13.31}$$

where K is the gain of the difference detector.

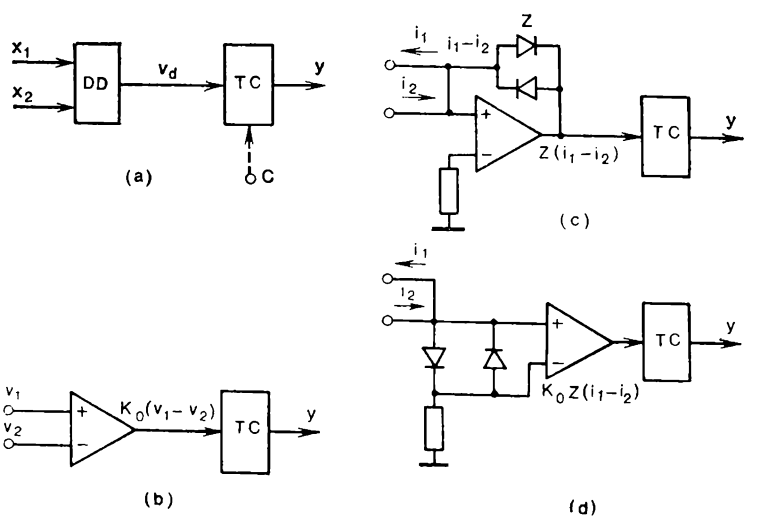


Fig. 13.51

The threshold circuit converts the analog signal v_d into a binary variable. The type of threshold circuit used determines the functional capabilities of the entire comparator. The threshold circuit may be treated as a zero-crossing detector (the reference is set to zero), for which the systematic, dynamic and random errors are precisely known. For the threshold circuit, the error of primary importance is the systematic comparison error in the threshold, critical or triggering voltage. In our further discussion, the threshold voltage will be designated V_{th}, and the input current associated with the threshold voltage will be denoted I_{th}.

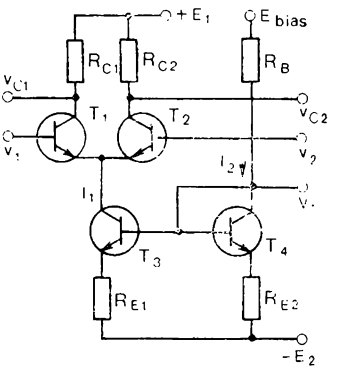


Fig. 13.52

Consider the principles underlying the circuit realizations of an amplitude comparator.

Difference detector. In voltage comparators (Fig. 13.51b), as in current comparators (Fig. 13.51*c* and *d*), the function of the difference detector, in accord with Eq. (13.31), is performed by a differential opamp with gain K_0. In current comparators, the difference signal is converted into a voltage across the nonlinear impedance Z of two parallel-connected diodes. Among the advantages offered by the circuit in Fig. 13.51*c* is low input resistance and freedom from the limiting action in the amplifier. The circuit shown in Fig. 13.51*d* has a better stability and has therefore a higher gain. The nonlinearity of Z limits the voltage amplitude and protects the amplifier against overloading.

Most frequently, comparators are based on IC differential opamps built as physically and functionally self-contained packages. The circuit of a typical stage in such an opamp is shown in Fig. 13.52. Its operation is based on the assumption that all the transistors have the same characteristics. This assumption is valid because the components are manufactured by IC technology. The circuit consists of a current switch using transistors $T1$ and $T2$ (see Sec. 5.5) and a current generator using transistors $T3$ and $T4$. Assuming in Eq. (5.52) that $I=I_1$, $M=1$, $v_x - E_B = v_1 - v_2 = v_d$, the transfer characteristic of the differential stage may be written

$$v_{C1} = E_1 - \frac{I_1 \alpha R_{C1}}{1+\exp(-v_d/m\varphi_T)}$$
$$v_{C2} = E_1 - \frac{I_1 \alpha R_{C2}}{1+\exp(v_d/m\varphi_T)} \qquad (13.32)$$

The gains K_1 and K_2 are found as the derivatives dv_{C_1}/dv_d and dv_{C_2}/dv_d at point $v_d = 0$. On differentiating Eq. (13.32), we get

$$K_1 = -I_1 \alpha R_{C_1}/4m\varphi_T$$
$$K_2 = I_1 \alpha R_{C_2}/4m\varphi_T \qquad (13.33)$$

In practice, symmetrical stages are used, in which $R_{C_1} = R_{C_2} = R_C$. Then $-K_1 = +K_2 = I_1 \alpha R_C/4m\varphi_T$. Hence, it is seen that the gain of the differential stage is a function of the maximum amplitude of the change in the output signal, equal to $I_1 \alpha R_C$. The value of $I_1 \alpha R_C$ is limited by the requirement for the circuit to be able to handle a wide range of changes in the input voltages. Therefore, the stage gain does not usually exceed 100. If a higher gain is desired, recourse may be had to multistage amplifiers. If the next amplifier stage has a differential input, the output voltage will be $v_{C_2} - v_{C_1}$. If so, the stage gain is doubled.

In the simple amplifier stage examined above, the transistors can be made to operate in the active region by shifting the range of changes in output voltage into the positive region. In practice, one usually includes suitable feedback shifting elements so that a bipolar signal can be obtained at the opamp output (Fig. 13.53). At first glance, it would seem that a better choice would be to use a differential stage based on *pnp*-transistors (Fig. 13.53*a*), because it can both shift the output voltage and amplify the signal. However, IC silicon *pnp*-transistors have a low speed.

In contrast, a higher speed is obtained with a shifting circuit containing a Zener diode, *D* (Fig. 13.53*b*). It is employed when the output voltage is to be shifted by a fixed amount. High speed can alternatively be obtained with a common-base transistor circuit (Fig. 13.53*c*), because the r.f. component of the signal has its path

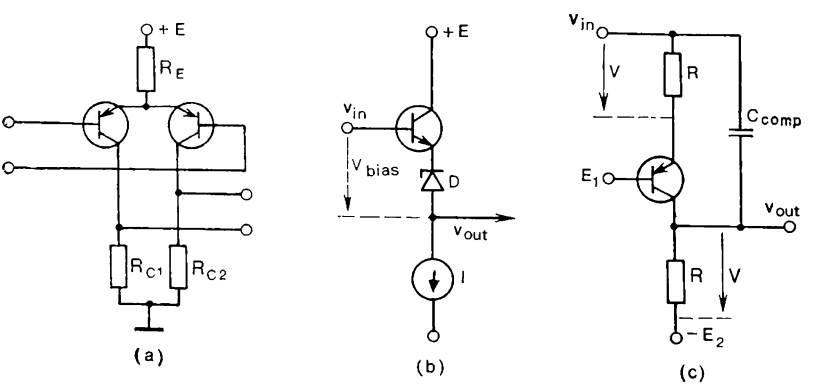


Fig. 13.53

via a compensating capacitor, C_{comp}. An advantage of this arrangement is that the amount of shift can be adjusted at will by varying E_1 or E_2.

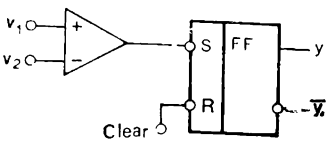


Fig. 13.54

Threshold circuit. It can be made with flip-flops, multivibrators, blocking oscillators, or tunnel-diodes. The more recent trend stimulated by advances in digital IC technology has been to use flip-flops based on level-type logic gates. The triggering threshold of such a circuit is determined by the value of V_{th} on the static transfer characteristic of the logic-gate network which carries the signal as far as the point of closure of the feedback loop. Variations in the triggering threshold of threshold circuits based on general-purpose logic gates may run into hundreds of millivolts. High accuracy of comparison can be achieved using either a high-gain difference detector or automatic error compensation [25].

Asynchronous threshold circuits are made with asynchronous flip-flops. Most often, this is a circuit using RS flip-flops (Fig. 13.54). Initially, the flip-flop is reset by applying a "Clear" signal. The flip-flop changes to a 1 state in response to the leading edge of the signal from the opamp, which is formed just as the difference $v_1 - v_2$ crosses zero. This is a unidirectional circuit. A comparator capable of operating in both directions can be obtained by using two amplifiers, one supplying the signal to set the flip-flop to a 1 state, and the other a signal resetting the flip-flop to a 0 state. The RS flip-flop shown in Fig. 13.55*a* goes to a 1 state just as the positive-going voltage $v(t)$ crosses the upper threshold level v_1 (Fig. 13.55*b*), and to a 0 state just as the negative-going voltage $v(t)$ crosses the lower threshold level v_4. The combination $SR = 1$ is forbidden; its appearance at the input of the flip-flop is prevented by satisfying the condition $v_1 > v_4$. This circuit configuration may also be used when the input voltages v_2 and v_3 change independently of each other. In this case, the conditions $v_2 > v_1$ and $v_4 > v_3$ must not appear simultaneously.

No restrictions are imposed on changes in the input signals for a comparator in which the threshold circuit is a JK flip-flop (Fig. 13.56*a*). The manner in which such a comparator will operate depends on which inputs accept $v(t)$. Figure 13.56 shows the case where $v(t)$ is applied to the inverting inputs of both opamps. With this arrangement, the flip-flop is switched to a 1 and a 0 state by the positive-going edge of $v(t)$ (Fig. 13.56*b*). Because $v(t)$ is a linear function, the pulse appearing at the comparator output has a duration proportional to the difference $v_4 - v_1$.

As a rule, synchronous threshold circuits are made with clocked D flip-flops. The synchronous threshold circuit shown in Fig. 13.57*a*

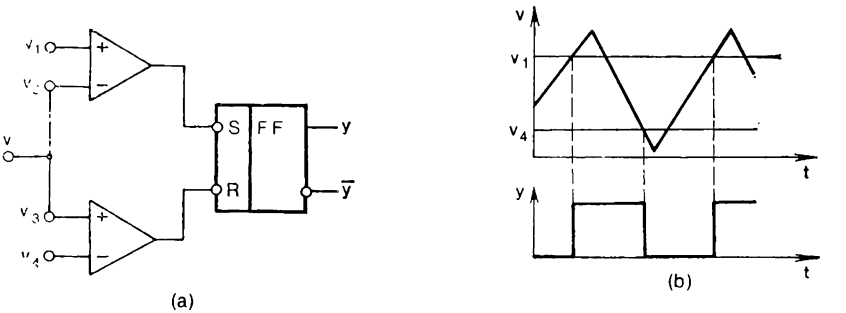


Fig. 13.55

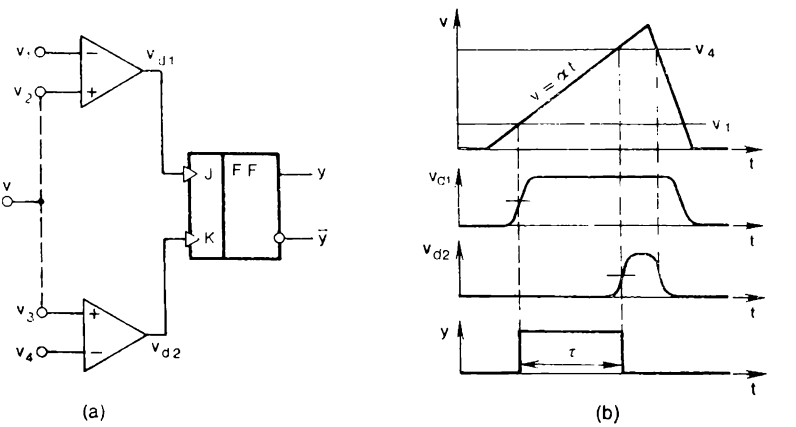


Fig. 13.56

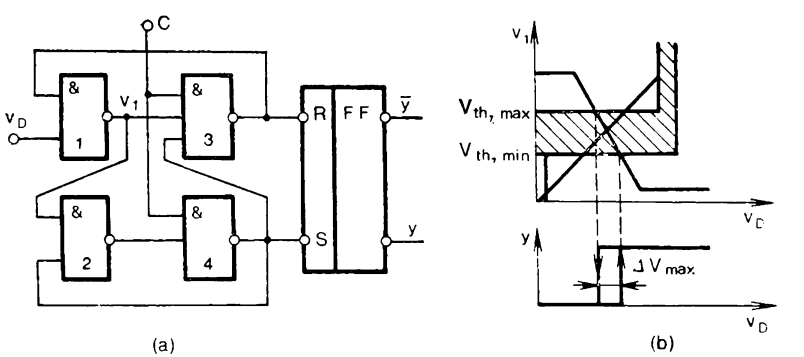


Fig. 13.57

is essentially a lockout D flip-flop. It would have an ideal threshold characteristic, if the logic gates forming its control section were identical in performance. The point is that, because there would be no uncertainty region, any input voltage would be recognized by the circuit as either a 1 or a 0. Therefore, after the application of a clock pulse enabling gates *3* and *4*, one of them would go into a 0 state, and the associated low-level voltage would actuate the lockout feature. In practice, however, the component logic gates differ in their threshold levels, so it may well happen, at least in principle, that v_1 at the output of gate *1* may be recognized by gates *3* and *4* as a logic 0 and by gates *2* and *4* as a logic 1. Then, the application of a clock pulse would leave gates *3* and *4* in the 1 state, the lockout feature will not operate, and the flip-flop cell will remain in the previous state. The maximum width of the uncertainty region, ΔV_{max}(Fig.13.57*b*) is equal to the ratio of the uncertainty region width, $V_{th,\,max} - V_{th,\,min}$, to the gain of the input logic gate. Because of this, it is a fraction of the likely instability in the threshold level. If the associated difference detector has a high gain, the uncertainty region of the D flip-flop will have practically no effect on the performance of the comparator owing to the existence of noise voltages.

In a latch D flip-flop, no uncertainty region can exist for basic reasons.

Phase Comparator

In a phase comparator, the input signal is the difference in phase between two pulse trains, x_1 and x_2, shown in Fig. 13.58*a*. The phase is measured in terms of the time interval between the positive-going edges of the pulses.As is seen from Fig. 13.58*a*, if the phase difference is positive ($\varphi_2 - \varphi_1 > 0$) during the positive-going edge of $x_2\,(t)$, the variable x_1 is a logic 1; if the phase difference is negative ($\varphi_2 - \varphi_1 < 0$), it is a logic 0. Therefore, the function of a phase comparator can be performed by a clocked D flip-flop (Fig. 13.58*b*).

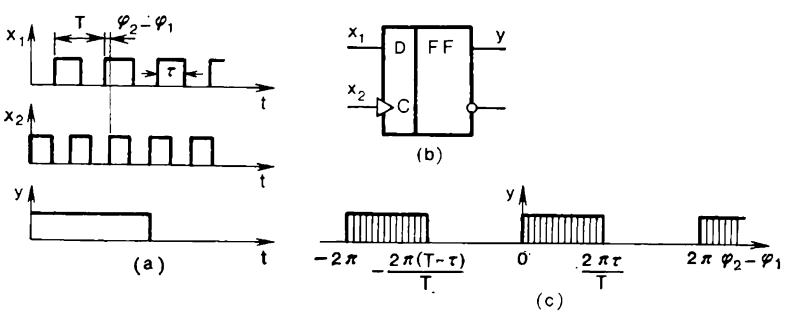


Fig. 13.58

Periodicity in the transfer characteristic $y(\varphi_2 - \varphi_1)$ of a phase comparator (Fig. 13.58*c*) arises from the fact that the phase difference can be determined accurate to the term $2\pi n$ (where n is any integer). The error in operation of a phase comparator has its source in the propagation delay through the logic gates of the flip-flop. Therefore, the resolution of a phase comparator is improved as the speed of the flip-flop it uses is increased. The systematic comparison error is the phase difference such that the flip-flop can go to a 0 or a 1 state with an equal probability.

Frequency Comparator

The output of a frequency comparator is a binary signal y whose value is determined by the sign of the difference in frequency, $f_1 - f_2$, between two periodic pulse trains, x_1 and x_2. Basically, the comparison of frequencies f_1 and f_2 consists in detecting and analyzing the instants at which the leading edges of pulses in trains x_1 and x_2 occur (in the timing diagrams, these instants are numbered consecutively). Like a phase comparator, a frequency comparator detects two equiprobable conditions (provided it is free from a systematic error), namely whether x_1 occurs ahead of x_2, or vice versa. So, the joint occurrence of x_1 and x_2 is left out of consideration.

If $f_1 = f_2$, x_1 and x_2 occur alternately with precision. If $f_1 \neq f_2$, this regularity is violated. To make our discussion more definite, suppose that $f_1 > f_2$. Let τ_2 denote the time interval within which x_2 occurs by unity more frequently than x_1. On writing

$$f_2\tau_2 - f_1\tau_1 = 1$$

we find that

$$\tau_2 = 1/(f_2 - f_1) \tag{13.34}$$

An arbitrary τ_2 can contain only one pair of adjacent x_1's (they are marked t_5 and t_8 in Fig. 13.59*b*) between which two x_2's can occur

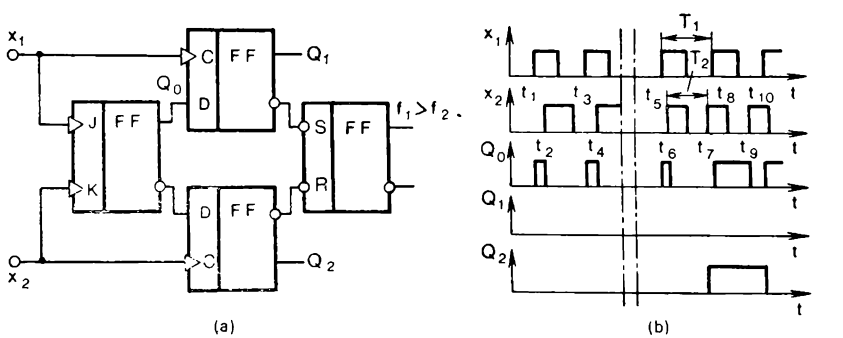


Fig. 13.59

(the respective occurrence times are t_6 and t_7). Similarly, if $f_1 < f_2$, the time interval $\tau_1 = 1/(f_1 - f_2)$ will contain only one pair of adjacent x_2's between which two x_1's can occur. Such a departure from an alternating occurrence of x_1 and x_2 is repeated at a frequency $f_d = |f_1 - f_2|$.

A frequency comparator is an asynchronous sequential network with two true dynamic inputs x_1 and x_2. Its algorithm may be stated as follows.

1. The output state remains unchanged, if the inputs accept a train of precisely alternating signals x_1 and x_2.
2. The output goes to a 1 state, if the input x_1 accepts a contiguous pair of signals not separated by an x_2 signal.
3. The output goes to a 0 state, if a contiguous pair of signals is applied to input x_2.

The circuit of a frequency comparator operating in the above sequence is shown in Fig. 13.59*a*, and its timing diagrams in Fig. 13.59*b*.

If the trains of input pulses are interleaved regularly, every next pulse causes the asynchronous JK flip-flop to change state in such a way that when an x_1 is applied the flip-flop is in the state $Q_0 = 0$, and when an x_2 is applied, it is in the state $Q_0 = 1$ always.

Should two x_1 signals come contiguously, the JK flip-flop will be in a 1 state by the time the second x_2 arrives. As a result, the second x_1 signal will write a 1 into the upper D flip-flop, $Q_1 = 1$, and it will set the output RS flip-flop to a 1 state.

Similarly, should two x_2 signals come contiguously, the JK flip-flop will be in a 0 state by the time the second x_2 arrives. As a result, a 1 will be written into the lower D flip-flop, and the signal $Q_2 = 1$ will switch the output flip-flop to a 0 state.

Pulses occur at output Q_1 of the upper flip-flop when $f_1 > f_2$ and they do so at a frequency $f_d = f_1 - f_2$. Pulses occur at output Q_2 of the lower flip-flop when $f_2 > f_1$, and they do so at a frequency $f_d = f_2 - f_1$.

The above comparator uses no clock pulses, so it is an asynchronous comparator.

The output of a frequency comparator goes from 0 to 1 or back with a delay relative to the instant when the difference signal $f_1 - f_2$ changes sign. In accord with Eq. (13.34) the maximum delay is given by

$$\tau_{d,\max} = 1/|f_1 - f_2|$$

The mean delay is a half of the maximum one.

Consider the case where the difference signal varies linearly, $f_1 - f_2 = At$. The time interval t_1 which elapses since the difference $f_1 - f_2$ changes sign and which contains a number of x_2 pulses by unity greater than the number of x_1 pulses can be found, using

the following equation

$$\int_0^{t_1} At\,dt = 0.5At_1^2 = 1$$

Hence, $t_1 = \sqrt{2A^{-1}}$. The time interval t_1 defines the maximum delay with which the comparator changes state. The mean delay can be found in a similar way, using the equation

$$0.5A\tau_d^2 = 0.5$$

Hence,

$$\overline{\tau_d} = \sqrt{1/A}$$

The systematic error referred to the comparator input will be

$$\overline{(f_1 - f_2)} = A\overline{\tau_d} = \sqrt{A}$$

The systematic error decreases with decreasing rate of change of the difference signal.

13.6. PULSE TIME GATES

The function of a pulse time gate is to reproduce an incoming signal at its output for the duration of a definite (gating) time interval. Outside the gating interval, there should be no output signal. The gating interval is determined by a control signal ordinary called a *gating pulse*, usually rectangular in shape. A particular form of the pulse time gate is the AND gate (see Sec. 6.6) which will block or pass a signal, depending on the control (enable or disable) signal applied. We shall discuss circuits intended to gate analog pulse signals.

Pulse time gates may be classed into the series type (Fig. 13.60*a*), the parallel type (Fig. 13.60*b*), the series-parallel type (Fig. 13.60*c*), and the current type (Fig. 13.60*d*), depending on how the switch is connected in the circuit. The first three are fully analogous in terms of merits and demerits to the series, parallel, and series-parallel amplitude limiters examined in Chap. 4. The main requirement that their switches are to meet is to satisfy the inequality $r_{ON} \ll R \ll r_{OFF}$. In contrast, in the circuit of Fig. 13.60*d* the first of these

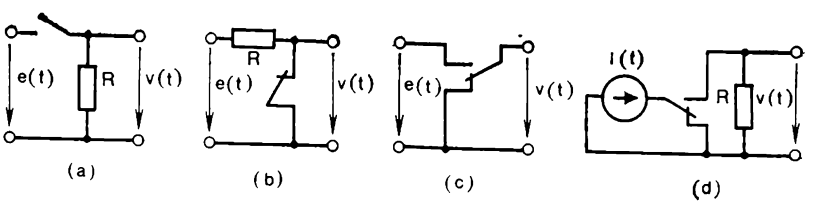


Fig. 13.60

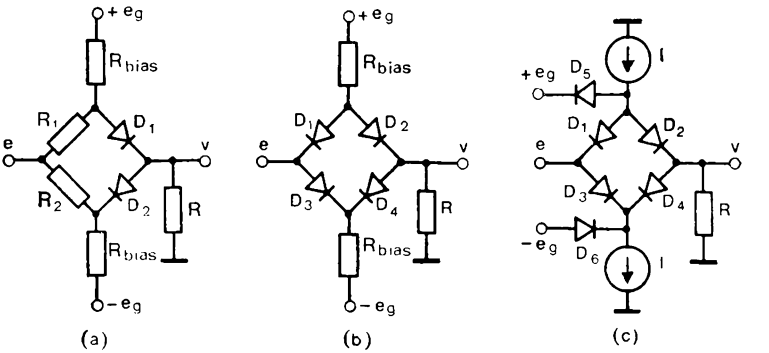


Fig. 13.61

inequalities has practically no effect on the accuracy of the pulse time gate.

Pulse time gates may use transistor and diode switches. Consider three basic types of the bridge diode switch (Fig. 13.61) employed in series pulse time gates.

The switch is controlled by a pair of complementary signals, e_g and $-e_g$. For the switch to be closed, it is required that $e_g > e_0$, in which case all the diodes in the bridge will be conducting. For the correct operation of the circuit it is further required that the amplitude of the signal $e\ (t)$ be such that

$$|e|_{max} < |e_g| + e_0$$

A limitation of the two-diode circuit (Fig. 13.61*a*) is a relatively high resistance in the ON state: $r_{ON} \approx R_1 = R_2$. In the circuit of Fig. 13.61*b* and *c*, the ON resistance is solely decided by the forward resistance of the diodes, $r_{ON} \approx 2r_f$. Among the advantages of the circuit in Fig. 13.61*c* is that R_{bias} do not shunt the load resistor R, so that the effect of the tinternal source resistance is reduced. Similar diode switches may be used in parallel pulse time gates.

Referring to the current time gate (Fig. 13.62), its state is determined by that of the current switch using transistors $T1$ and $T2$. In the ON state, transistor $T2$ is conducting. Now practically all of the emitter current, $i_E\ (t)$, of $T3$, determined by the input signal $e\ (t)$, is flowing

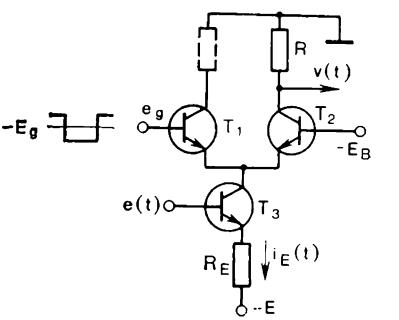


Fig. 13.62

through the load resistor R, thereby producing the output signal $v(t)$. In the OFF state, transistor $T2$ is cut off, and $v(t) = 0$.

13.7. VARIABLE PULSE DELAY CIRCUITS

A variable pulse delay circuit generates an output pulse which appears at a time delayed for a time τ_d from the appearance of a trigger pulse, v_{trig}. The value of τ_d is determined by the control signal which may be an analog quantity (say, a control voltage e_c) or a digital quantity (say, a numerical code N_c). In most cases, it is required that τ_d be a linear function of e_c or N_c.

A typical block diagram of an analog delay circuit and its timing diagrams are shown in Fig. 13.63. The initial state of the circuit is determined by that of the flip-flop, $Q = 0$. The leading edge of a trigger pulse, v_{trig}, sets the flip-flop to a 1 state and triggers the ramp generator. Taking the trigger pulse as the origin of time, the manner in which the voltage at the generator output varies may be

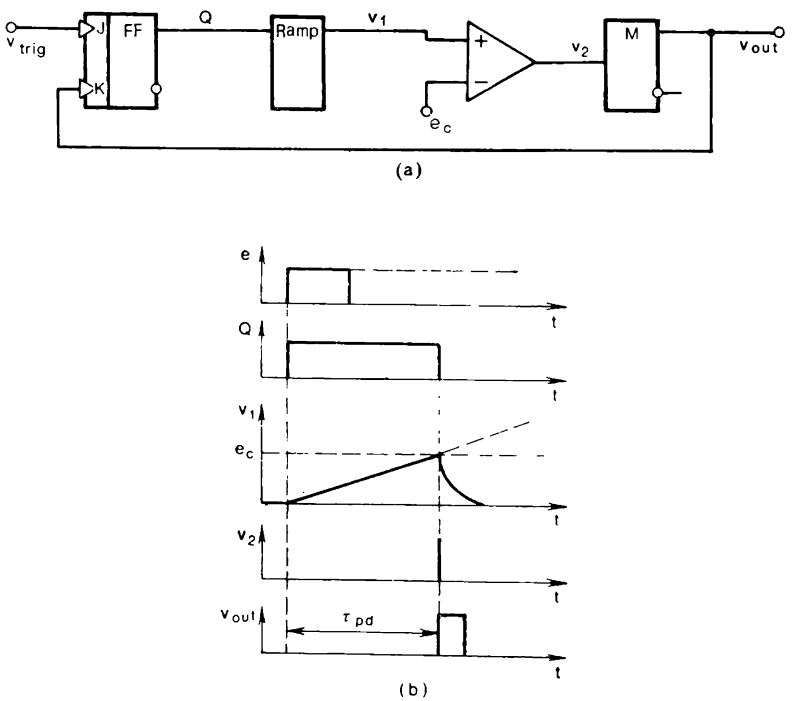


Fig. 13.63

written as $v_1(t) = At$. At the instant when

$$At = e_c \tag{13.35}$$

the comparator is switched and the one-shot multivibrator, M, is triggered. The leading edge of the pulse appearing at the multivibrator output resets the flip-flop, thereby switching the circuit back to its initial state. Solving Eq. (13.35) for time, we find that the output pulse of the multivibrator is delayed from the leading edge of the trigger pulse by a time interval

$$\tau_d = e_c/A \tag{13.36}$$

In real delay circuits, Eq.(13.36) is satisfied only approximately because of the errors introduced by the ramp generator and comparator. The per cent error may be reduced to less than 1% through the use of precision components.

The above delay circuit uses discrete elements. If the requirements for the accuracy of time delay are very stringent, the complete equipment may be simplified by combining the ramp generator, comparator, and flip-flop. An example is the phantastron generator examined in Chap. 9.

In digital pulse delay circuits, $v_1(t)$ and e_c are respectively replaced by the numbers N and N_c. The ramp generator is replaced by a counter, and the analog comparator by a digital comparator (equality detector) (Fig. 13.64). Initially, $Q = 0$ and $N = 0$. Application of a trigger pulse sets the flip-flop to a 1 state, $Q = 1$. Following that, every next clock pulse will increment the number N by one. The flip-flop will reset to 0 immediately after the comparator detects the equality $N = N_c$. The time delay is given by $\tau_d = T_{sync}N_c$, where T_{sync} is the repetition period of sync (clock) pulses.

The per cent error of digital pulse delay circuits depends on the instability in T_{sync}. Through the use of crystal-controlled clock oscil-

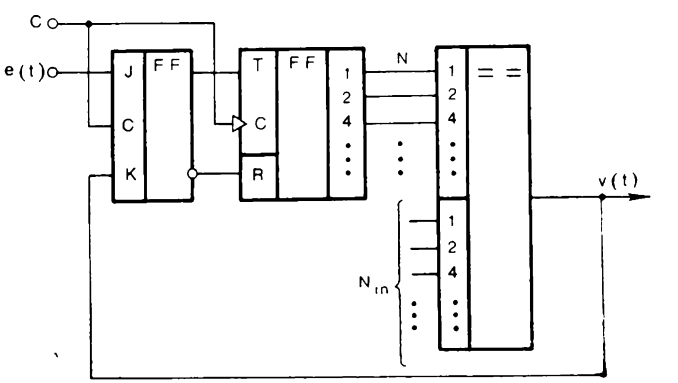


Fig. 13.64

lators, the error may be reduced to anywhere between 10^{-4} and 10^{-5}, practically unattainable with analog delay circuits.

Table 13.3. Sequence of Control Signals

t	R'	V_1	V_2
0	0	0	0
1	1	0	0
2	1	1	0
3	1	0	0
4	1	0	1
5	1	0	0
6	0	0	0

Table 13.4. Sequence of Input Signals

	Vt_{WT}	$\overline{V}_S\, dV_S$	$\overline{R}$	$\overline{S}$	Sgn	D_R	R_L
0	0	0	0	1	0	1	0
1	1	0	1	1	0	1	1
2	0	1	1	1	0	1	0
3	0	1	1	1	0	1	0
4	0	0	1	1	1	1	0
5	0	1	1	1	1	1	0
6	0	1	1	1	1	0	1
7	0	0	1	1	1	0	1
8	0	0	0	1	1	1	0

REVIEW QUESTIONS AND PROBLEMS

13.1. What sets serial static registers apart from parallel ones? Compare them in terms of the number of control inputs used, word read-in and read-out, and the number of input and output terminals.

13.2. Point out the merits and demerits of dynamic registers as compared with static registers. Draw a block diagram for a parallel dynamic register.

13.3. Input A_1 of an 8-bit parallel register (see Fig. 13.1) accepts an input word 11000100, whereas input A_2 of the same register accepts another word, 01010101. Using the sequence of control signals exhausted in Table 13.3, find the word stored in the register.

13.4. Assuming that the discrete time interval t is known during the existence of the clock pulse at input C and neglecting the value of R', solve Problem 13.3 for the registers arranged as shown in Figs. 13.2 and 13.3.

13.5. A three-bit serial register is built as shown in Fig. 13.15. Using the input signals exhausted in Table 13.4, find the sequence of signals at outputs Y_R and Y_L, if the input word applied to its parallel inputs D_1 through D_3 is 100.

13.6. Draw a logic diagram for one stage of the register shown symbolically in Fig. 13.6*a*.

13.7. Draw logic diagrams for one stage of a serial-carry binary counter (Fig. 13.24*b*) using NAND gates and AND-NOR gates. Compare the two circuits for speed, assuming that $\bar{\tau}_d$ is the same for both.

13.8. Draw a block diagram for one stage of an up-counter and an up-down counter.

13.9. Outline ways and means for enhancing the speed of counters.

13.10. Draw a logic diagram for a NAND-gate three-bit asynchronous parallel-carry up-counter. Evaluate the dynamic parameters of the counter.

13.11. Draw a circuit diagram for an analog variable delay circuit, using Soviet-made 1УТ401Б opamps and Series 133 logic gates (NAND and AND-NOR), and also the flip-flops shown in Figs. 6.31, 6.38, and 6.39*a*.

13.12. Design a digital functional subunit generating a 1 signal, if the binary variable x_1 takes on consecutive values such that $x_1(t) = 0$, $x_1(t-1) = 1$, $x_1(t-5) = 1$, and $x_1(t-6) = 0$. The discrete time intervals are marked by the positive-going edge of another binary variable, x_2. Hint: use a serial-parallel register and a code comparator.

13.13. Design a functional subunit differing from that specified in Problems 13.12 in that the appearance of the given sequence of values switches the output variable from 0 to 1. It is switched back from 1 to 0 by a sequence of signals such that $x_1(t) = 1$, $x_1(t-2) = 1$, $x_1(t-3) = 0$, and $x_1(t-6) = 1$.

13.14. In a five-bit serial register, output $\bar{Y}_R$ is connected to input D_R. Enable input V_R to accept a periodic sequence of pulses. Find the values of the word stored in the register at the successive time instants $t = 1, 2, \ldots$, if at $t = 0$ it is 00000.

13.15. Using the values of the word found in Problem 13.13 to code the natural numbers from 0 through 9, design a modulo-10, modulo-100, and modulo-10^n asynchronous counter.

BIBLIOGRAPHY

1. Меерович Л. А., Зеличенко Л. Г., *Импульсная техника.* 2-е изд., Сов. радио, Москва, 1954.
2. Ицхоки Я. С., *Импульсные устройства.* Сов. радио, Москва, 1959.
3. Гольденберг Л. М., *Импульсные и цифровые устройства.* Связь, Москва, 1973.
4. Фролкин В. Т., *Импульсные устройства.* Машиностроение, Москва, 1966.
5. Миллман Я., Тауб Г., *Импульсные и цифровые устройства.* Пер. с англ. под ред. Л. А. Мееровича, Госэнергоиздат, Москва, 1960.
6. Айзинов М. М., *Переходные процессы в элементах радиоустройств.* Морской транспорт, Ленинград, 1955.
7. Гоноровский Г. К., *Радиотехнические цепи и сигналы.* Сов. радио, Москва, 1977.
8. Гаврилов Г. К., *Приближенные методы анализа переходных процессов.* Сов. радио, Москва, 1966.
9. Литвиненко О. Н., Сошников В. И., *Теория неоднородных линий и их применение в радиотехнике.* Сов. радио, Москва, 1964.
10. Литвиненко О. Н., Сошников В. И., *Расчет формирующих линий.* Гостехиздат УССР, Киев, 1962.
11. Лукин Ф. В., *Переходные процессы в линейных элементах радиотехнических устройств.* Оборонгиз, Москва, 1960.
12. Важенина З. П., Волкова Н. Н., Чадович И. И., *Методы и схемы временной задержки импульсных сигналов.* Сов. радио, Москва, 1971.
13. Меерович Л. А., Ватин И. М., Зайцев Э. Ф., Кандыкин В. М., *Магнитные генераторы импульсов.* Сов. радио, Москва, 1968.
14. Степаненко И. П., *Основы теории транзисторов и транзисторных схем.* Энергия, Москва, 1977.
15. Шац С. Я., *Транзисторы в импульсной технике.* Судпромгиз, Ленинград, 1963.
16. Агаханян Т. М., *Электронные ключи и нелинейные импульсные усилители.* Сов. радио, Москва, 1966.
17. Яковлев В. Н., *Импульсные генераторы на транзисторах.* Техника, Киев, 1968.
18. Доронкин Е. Ф., Воскресенский В. В., *Транзисторные генераторы импульсов.* Связь, Москва, 1968.
19. Гусев В. П., Фомин Г. М., и др., *Расчет электрических допусков радиоэлектронной аппаратуры.* Сов. радио, Москва, 1963.
20. Ильин В. А., Коншин Б. Н., *Импульсные устройства с мостовыми времязадающими цепями.* Энергия, Москва, 1972.
21. Моругин Л. А., *Импульсные схемы на туннельных диодах.* Сов. радио, Москва, 1966.
22. Горяинов С. А., Абергауз И. Д., *Полупроводниковые приборы с отрицательным сопротивлением.* Энергия, Москва, 1970.
23. Важенина З. П., *Фантастронные генераторы. Теория проектирования, расчет.* Сов. радио, Москва, 1968.
24. Наумов Ю. Е., *Интегральные логические схемы.* Сов. радио, Москва, 1970.

25. Гитис Э. И. *Преобразователи информации для электронных цифровых вычислительных устройств.* Госэнергоиздат, Москва, 1975.
26. Алексеенко А. Г.,*Современная микросхемотехника.* Энергия, Москва, 1979.
27. Шац С. А. (ред.), *Проектирование радиоэлектронных устройств на интегральных микросхемах.* Сов. радио, Москва. 1976.
28. Ильин В. Н., *Основы автоматизации схемотехнического проектирования.* Энергия, Москва, 1979.
29. Трохименко Я. К., Каширский И. С., Ловкий В. К., *Проектирование радиотехнических схем на инженерных ЭЦВМ.* Техника, Киев, 1976.
30. Агаханян Т. М., Гаврилов Л. Е., Мищенко Б. Г., *Основы наносекундной импульсной техники.* Атомиздат, Москва, 1976.
31. Ицхоки Я. С., Овчинников Н. И., *Импульсные и цифровые устройства.* Сов. радио, Москва, 1972.
32. Ильин В. Н., Фролкин В. Т., *Цифровые схемы и устройства на МДП транзисторах.* Энергия, Москва, 1975.
33. Берман В. Р., Фролкин В. Т., *Генераторы магнитной развертки электронно-лучевых индикаторов.* Сов. радио, Москва, 1976.
34. Носов Ю. Р., Петросьянц К. О., Шилин В. А., *Математические модели элементов интегральной электроники.* Сов. радио, Москва, 1976.
35. Гаврилов Г. К., *Переходные процессы в транзисторе и методы расчета импульсных схем.* Связь, Москва, 1971.
36. Яковлев В. Н. (ред.) *Справочник по импульсной технике.* Техника, Киев, 1970.
37. Букреев И. Н., Мансуров Б. М., Горячев В. И., *Микроэлектронные схемы цифровых устройств.* Сов. радио, Москва, 1975.
38. Казаринов Ю. М. (ред.), *Расчет импульсных и цифровых схем радиотехнических устройств.* Высшая школа, Москва, 1976.
39. Агаханян Т. М. ред.), *Расчет импульсных устройств на полупроводниковых приборах. Сборник примеров и задач.* Сов. радио, Москва, 1975.
40. Гольденберг Л. М. (ред.), *Расчет и проектирование импульсных устройств* Сов. радио, Москва, 1975.
41. Millman J., Taub H., *Pulse Digital and Switching Waveforms*, McGraw-Hill Book Co., New York, N.Y., 1970.
42. Millman J., Halkias Ch., *Integrated Electronics: Analog and Digital Circuits and Systems*, McGraw-Hill Book Co., New York, N. Y., 1972.
43. Аваев Н. А., Дулин В. Н., Наумов Ю. Е., *Большие интегральные схемы с инжекционным питанием.* Сов. радио, Москва, 1977.
44. Валиев К. А., Дягилев В. Н., и др., *Микромощные интегральные схемы.* Сов. радио, Москва, 1975.
45. Корж В. И., Шагурин И. И., «Графический метод определения помехоустойчивости цепочек логических элементов», *Изв. вузов. Приборостроение*, 1970, 13, **6**, 52-59.
46. Шагурин И. И., *Транзисторно-транзисторные логические схемы.* Сов. радио, Москва, 1974.
47. Lueke, Gerald,"Noise margins in digital integrated circuits", *Proc. IEEE*, 1964, 52, **12**, 1565.
48. Gill Art., *Introduction to the Theory of Finite-State Machines*, McGraw-Hill Book Co., New York, N. Y., 1961.
49. Миллер Р., *Теория переключательных схем.* В 2-х т. Перев. с англ. под ред. П. П. Пархоменко, Наука, Москва, 1971.
50. Майоров С. А. (ред.), *Проектирование микроэлектронных устройств.* Сов. радио, Москва, 1977.
51. Угрюмов Е. П., *Элементы и узлы ЭЦВМ.* Высшая школа, Москва, 1976.
52. Майоров С. А. (ред.), *Проектирование цифровых вычислительных машин.* Высшая школа, Москва, 1972.
53. Dulin V., *Electron Devices*, Mir Publishers, Moscow, 1980.

54. Ерофеев Ю. Н., *Основы импульсной техники*. Высшая школа, Москва, 1979.
55. Якубовский С. В. (ред.), *Аналоговые и цифровые интегральные схемы*. Сов. радио, Москва, 1979.
56. Носов Ю. Р., *Полупроводниковые диоды с накоплением заряда и их применение*. Сов. радио, Москва, 1966.

INDEX

A

Active region, 72
Adder, 359
  full, 359
  half, 359
Addition
  parallel, 359
  serial, 359
Analog comparator, 357
Analog detector, 367
AND-NOR gate, 145
AND operation, 115

B

Basic set in switching algebra, 115
Binary cell, 352
Binary coding
  level type, 11
  pulse type, 11
Binary equality detector, 362
Bistable circuit, 161
  static, 161
  unsaturated, 177
  vacuum-tube, 177
Bit, 328
Blocking oscillator, 234
  collector-coupled, 234
  L-type, 246
  modified, 246
  vacuum-tube, 250
Bucket-brigade devices, 344

C

Capacitor, speed-up, 84
Charge-coupled devices, 344
Charge-transfer devices, 344
Clamping circuit, 107
Clear signal, 185
Clock pulse, 188
CML (ECL) gate, 146
Combinational logic, 180
Combinatorial logic, 180
Counter, 348
  binary, 351
  down-, 349
  reversible, 349
  ripple-through, 350
  up-, 348
  up-down, 349
Current boost, 305
Cutoff, dynamic, 72
Cutoff region, 72

D

Decoder, 363
Delay line, 37
  coaxial-cable, 41
  open-circuited, 39
  short-circuited, 39
  spiral, 42
  ultrasonic, 46
Demultiplexor, 363
Difference comparator, 376
Digital circuits, functional, 328
Digital equality detector, 362
Diode
  charge-storage, 50
  ideal, 51
  pulse, 50
  Schottky-barrier, 50
Displacement current, 34
Droop, 10
DTL gate, 137

F

Fan-in, 119
Fan-out, 119
Ferrites, 33
Finite automata, 179
Finite automata theory, 180
Finite-state machines, 178

Flip-flop, 161, 178
JK, 186
latch, 197
level logic gate type, 194
master-slave, 191
MS, 191
RS, 186
T, 186
thyristor, 325
tunnel-diode, 318
Flyback time, 252

G

Gating pulse, 378
Group velocity, 37

H

Horizontal scanning, 310

I

Inverter, 127
MOS-transistor, 155

L

L-oscillator, 246
L-section, 43
Ladder network, 43
Leading edge, 9
Leakage flux, 34
Limiter
common-base, 102
common-collector, 103
common-emitter, 100
diode, 94
double-peak, 93
emitter-coupled, 104
negative, 92
parallel, 94, 95
peak, 92
positive, 92
parallel, 94
positive-negative, 93
series, 95, 98
series-parallel, 95
transistor, 100
Logic gate
CMOS, 157
current-mode, 146
emitter-coupled, 146
level-type, 118
MOS-transistor, 153
MOSFET, 153
multi-input, 129
multistep, 137
pulse-type, 119

M

Magnetic hysteresis, 32
Master, 191
Microstrip lines, 43
MOS-transistor, 153
depletion-mode, 154
enhancement-mode, 154
n-channel, 154
p-channel, 154
MOSFET, 153
MOSFET inverter, 156
biased, 156
saturated-load, 156
unsaturated-load, 156
Multivibrator, 209
astable, 209
bistable, 161
emitter-coupled, 163
symmetrical, 167
unsymmetrical, 162
differential opamp as, 229
discrete-component, 210
free-running, 209
IC logic-gate type, 228
integrating *RC*-network, 229
IC logic-gate type, 222
integrated-circuit, 216
modified, 216
monostable, 209
collector-coupled, 210
one-shot, 209
integrating *RC*-network, 231
synchronized, 209
tunnel-diode, 316

N

Noise immunity, 124
Noise margin, 124
NOT operation, 115

O

Operational amplifier, 25
OR operation, 115
Overshoot, 9

P

Phantastron, 275
Phase comparator, 375
Phase velocity, 37
Pi-section, 43
Power-driver output stage, 143
Pulse circuits, 9
  analog, 10
  digital, 10
  functional, 328
  negative-resistance, 314
  tunnel-diode, 314
Pulse duration, 9
Pulse fall time, 9
Pulse length, 9
Pulse rate division, 289
Pulse rise time, 9
Pulse shaping, 14
Pulse shortening, 18
Pulse signals
  differentiation of, 24
  integration of, 24
Pulse time gate, 378
Pulse transformers, 30
Pulse width, 9

R

Ramp current, 293
Ramp generation, 256
Ramp generator, 252
  bootstrap, 266
  constant-current, 262
  CR, 257
  free-running, 256
  Miller-integrator, 273
  triggered, 256
Ramp voltage input, 23
Rectangular pulse input, 21
Register, 328
  dynamic, 341
  one-bit, 334
  parallel, 329
  parallel-serial, 340
  recirculating, 341
  serial, 331
  shift, 331
    bidirectional, 332
    unidirectional, 332
Relaxation oscillator
  avalanche-transistor, 319
  synchronization of, 285
  thyristor, 323
Reset signal, 185
Reset threshold, 163
Resonant circuit, 30
Response
  critically damped, 29
  overdamped, 29
  underdamped, 29
Ringing, 9

S

Sag, 10
Saturation region, 75
Satwooth current, 293
Sawtooth current generator, 293
  opamp-integrator, 307
  TV, 309
Sawtooth voltage, 252
Section
  constant-k, 44
  m-derived, 44
Sequential circuits, 178
  asynchronous, 182
  synchronous, 188
Set signal, 185
Set threshold, 163
Signal
  level-type, 11
  pulse-type, 11
Slave, 191
Slope error, 253
Square wave generator, 249
Step input, 21
Strip line, 43
Sweep generator
  negative-current-feedback, 299
  opamp-integrator, 307
Sweep time, 252
Sweep voltage, 252
Sweep-speed error, 253
Switch, 49
  diode, 50
    practical, 60
    transient response of, 61
  saturated, 261
  transistor, 63
    common-emitter, 69, 76
    modifications of, 83
  unsaturated, 86, 261
Switching algebra, basic set of, 115
Switching element, 49
Switching expression, 115
Switching frequency, 162
Switching functions, 114
Synchronization, 285

T

T-section, 43
Temperature stability, 69

Theory of finite-state machines, 180
Thyristor, 323
Tilt, 10
Time
  forward recovery, 56
  recovery, 162
  resolving, 162
  reverse recovery, 57
  storage, 57
  transition, 56, 57
  turn-on delay, 56
Time mark, 30
Time mark generator, 30
Timebase voltage, 252
Totem-pole amplifier, 143
Totem-pole output driver, 143
Trailing edge, 9
Transit time effects, 71
Transistor
  Ebers-Moll model, 64
  exponential model, 64
  MOS-153
  multi-emitter, 142
  Schottky, 87
  Schmitt, 316
Triggering
  asymmetrical, 169
  base, 169
  set-reset, 169
  steered, 177
  symmetrical, 169
Truth table, 114

V

Variable pulse delay circuit, 380
Vertical scanning, 309
Voltage boost, 305

W

Waveform, 9
  exponential, 9
  rectangular, 9
  sawtooth, 9
  trapezoidal, 9
  triangular, 9
Waveshape, 9
Word length, 328

Z

Zero-crossing detector, 368

Also from Mir Publishers. . .

AN INTRODUCTION TO COMPUTERS

By **N. Sergeev and N. Vashkevich**

Because of the rapid growth and wide use of computers, in practically all fields of man's activity, there is a need for students in all specialities to have a concise, yet thorough presentation of computer theory, design and operation.

This book is an attempt to fill this need. It covers the circuitry and operating principles of analog and digital computers, special-purpose computing devices, machines and systems. Ample space is devoted to a basic theory of similitude and simulation, and a fairly detailed description is given of basic functional elements, assemblies and units which make up typical analog and digital computers, in conjunction with an introduction to the mathematical and logical basis of electronic digital computers.

Separate sections deal with block-diagram synthesis of analog computers to solve algebraic, transcendental, ordinary and partial differential equations and their systems. The reader will find a survey of methos used to program problem solving on analog and digital computers, along with a brief outline of digital differential analyzers and hybrid systems.

Written by leading authorities in the field of Soviet computer engineering with a wealth of teaching and research experience at colleges, the book is student-oriented in the presentation of the subject-matter.

The book will be of primary value to college students and faculty members, and also to researchers and engineers concerned with the development and application of computers.

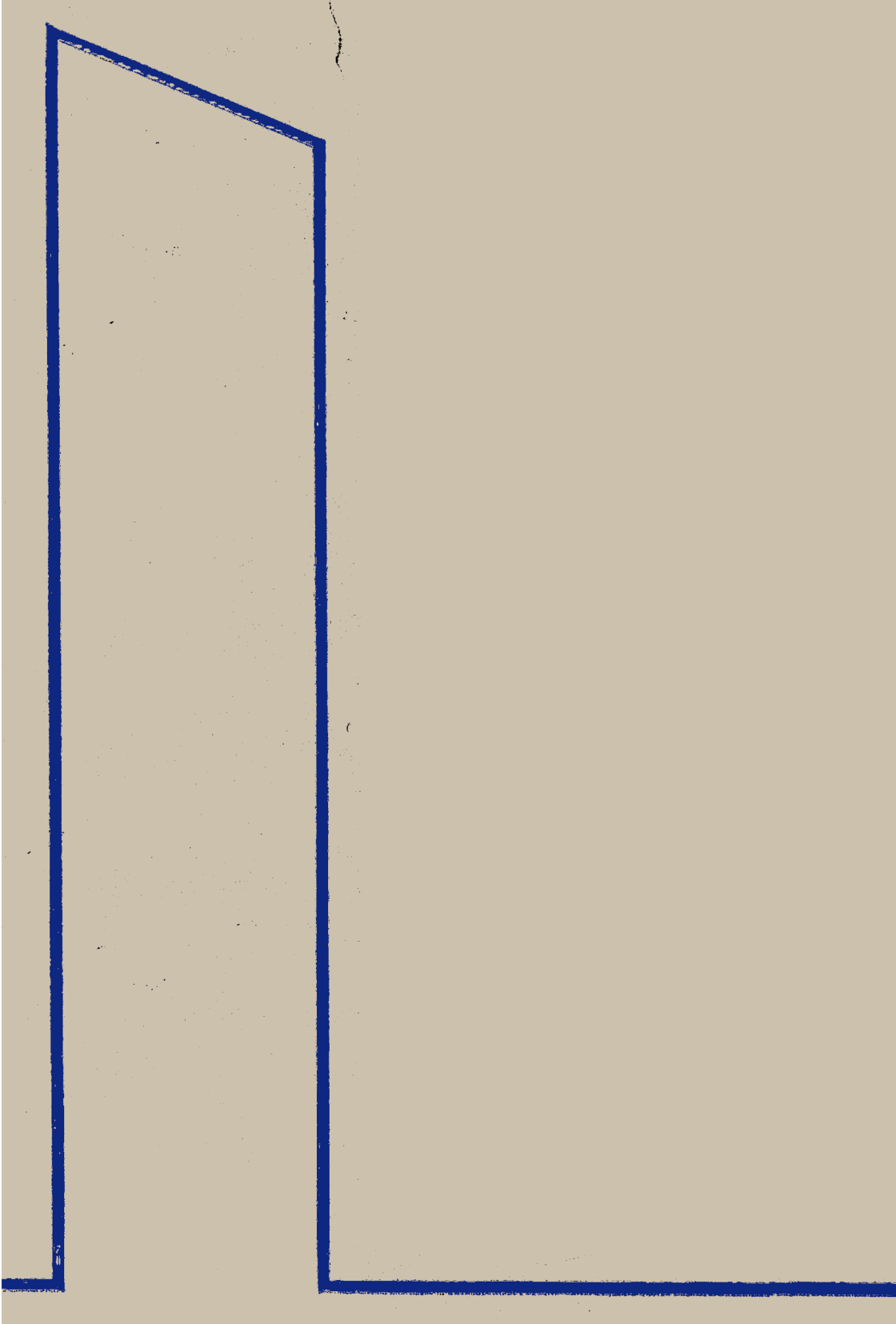
TO THE READER

Mir Publishers welcome your comments on the content, translation and design of this book.
We would also be pleased to receive any proposals you care to make about our future publications.

Our address is:
USSR, 129820, Moscow I-110, GSP
Pervy Rizhsky Pereulok, 2
Mir Publishers

Printed in the Union of Soviet Socialist Republics

ABOUT THE AUTHORS

Professor Viktor T. Frolkin, D.Sc. (Tech.), has been reading lectures on the subject for over 25 years at the Moscow Aviation Institute. He has penned over 50 publications, including monographs, texts, and study aids.

Docent Lev N. Popov, Cand.Sc. (Tech.), has been reading lectures on the subject for over five years at the Moscow Aviation Institute and also on pulse and digit microelectronics. Has written over 20 papers for technical journals and his department.